Wild Freitas da Silva Santos

Predição da Dinâmica de Lasers de Pontos Quânticos Semicondutores Utilizando Redes Neurais Artificiais

Editora Poisson

VOLUME 1

Wild Freitas da Silva Santos

Predição da Dinâmica de Lasers de Pontos Quânticos Semicondutores Utilizando Redes Neurais Artificiais

1ª Edição

Belo Horizonte
Editora Poisson
2023

Editor Chefe: Dr. Darly Fernando Andrade

Conselho Editorial
Dr. Antônio Artur de Souza – Universidade Federal de Minas Gerais
Ms. Davilson Eduardo Andrade
Dra. Elizângela de Jesus Oliveira – Universidade Federal do Amazonas
Msc. Fabiane dos Santos
Dr. José Eduardo Ferreira Lopes – Universidade Federal de Uberlândia
Dr. Otaviano Francisco Neves – Pontifícia Universidade Católica de Minas Gerais
Dr. Luiz Cláudio de Lima – Universidade FUMEC
Dr. Nelson Ferreira Filho – Faculdades Kennedy
Ms. Valdiney Alves de Oliveira – Universidade Federal de Uberlândia

Dados Internacionais de Catalogação na Publicação (CIP)

S237 Santos, Wild Freitas da Silva

Predição da Dinâmica de Lasers de Pontos Quânticos Semicondutores Utilizando Redes Neurais Artificiais/ Wild Freitas da Silva Santos – Belo Horizonte – MG: Editora Poisson, 2022

Formato: PDF
ISBN: 978-65-5866-241-9
DOI: 10.36229/978-65-5866-241-9
Modo de acesso: World Wide Web
Inclui bibliografia

1.Ensino 2. Educação I. SANTOS, Wild Freitas da Silva III. Título

CDD-620

Sônia Márcia Soares de Moura – CRB 6/1896

O conteúdo dos artigos e seus dados em sua forma, correção e confiabilidade são de responsabilidade exclusiva dos seus respectivos autores.

O conteúdo deste livro está licenciado sob a Licença de Atribuição Creative Commons 4.0.
Com ela é permitido compartilhar o livro, devendo ser dado o devido crédito, não podendo ser utilizado para fins comerciais e nem ser alterada.

www.poisson.com.br

contato@poisson.com.br

Dedicatória

Agradeço aos meus pais por terem me proporcionado a oportunidade de estudar e realizar os meus sonhos profissionais.

Agradecimentos

Ao Prof. Eduardo Simas, meus sinceros agradecimentos por toda a orientação e apoio recebido.

Ao Prof. George Thé meus agradecimentos por toda a coorientação, apoio recebido e possibilidade de estudar uma nova área na engenharia elétrica.

Aos professores John Butcher e Hugo Siqueira pelo suporte na implementação da solução ótima encontrada neste trabalho, fundamental para a finalização deste livro com sucesso.

A todos os familiares e amigos que contribuíram para a realização deste livro.

Apresentação

O livro, resultante do Mestrado em Engenharia Elétrica pela Universidade Federal da Bahia, discorre sobre a modelagem da dinâmica de lasers semicondutores de pontos quânticos (LPQ), uma nova classe de dispositivos eletrônicos com possibilidade de aplicações em telecomunicações e na medicina. O uso de LPQs requer atualmente a utilização de modelos tradicionais de equações de taxa ou em modelos de descrição de microestados, ambos são complexos e com elevados custos computacionais. Para a modelagem proposta neste livro, foram utilizadas amostras experimentais da potência óptica de saída do laser no estado fundamental (*Ground State* — GS) e no estado excitado (*Excited State* — ES) para temperaturas de operação de 20ºC, 40 ºC e 50 ºC. Redes Neurais Artificiais (RNA) como o Perceptron de Múltiplas Camadas (*Multilayer Perceptron* — MLP), Máquina de Aprendizado Extremo (*Extreme Learning Machine* — ELM) e Redes com Estados do Eco (*Echo State Network* — ESN) foram utilizadas como modelos de predição. Foi investigada também a aplicação das RNAs em conjunto com o Modelo Não-Linear Autoregressivo com Entradas Exógenas *(Nonlinear AutoRegressive models with eXogenous inputs* — NARX), onde as RNAs se apresentam como alternativa na estimação da não linearidade do modelo. Os resultados revelaram excelente previsibilidade com erros da ordem de 3,79% e um baixo custo computacional quando comparado às técnicas tradicionais.

Sumário

Dedicatória 4

Agradecimentos 5

Apresentação 6

1. INTRODUÇÃO 10

1.1. RELEVÂNCIA DO TEMA 10

1.2. OBJETIVO GERAL 11

1.3. OBJETIVOS ESPECÍFICOS 11

1.4. JUSTIFICATIVA 11

1.5. ORGANIZAÇÃO DA DISSERTAÇÃO 13

2. LASER SEMICONDUTORES DE PONTOS QUÂNTICOS 15

2.1. INTRODUÇÃO 15

2.2. DEFINIÇÕES E CARACTERÍSTICAS 15

2.3. APLICAÇÕES 21

2.4. MODELAGENS DO LASER DE PONTOS QUÂNTICOS 22

2.5. CONSIDERAÇÕES 23

3. MODELAGEM E PREDIÇÃO DE SÉRIES TEMPORAIS 24

3.1. INTRODUÇÃO 24

3.2. SÉRIES TEMPORAIS 24

3.3. A QUESTÃO FUNDAMENTAL: PREDIÇÃO 26

3.4. IDENTIFICAÇÃO DE SISTEMAS DINÂMICOS NÃO LINEARES 27

3.4.1. Modelo de Entrada 30

3.4.2. Sinal de Excitação 30

3.4.3. Arquitetura do Modelo 31

3.4.4. Representação Matemática da Dinâmica do Sistema 31

3.4.4.1. Modelo Autoregressivo Não Linear com Entradas Exógenas 35

3.4.5. Ordem do Modelo 36

3.4.6. Estrutura e Complexidade do Modelo 37

3.4.7. Parâmetros do Modelo 38

3.4.8. Validação do Modelo 38

3. 5. CONSIDERAÇÕES 39

4. INTRODUÇÃO ÀS REDES NEURAIS ARTIFICIAIS E AOS ALGORITMOS DE OTIMIZAÇÃO 40
4.1. INTRODUÇÃO 40
4.2. MODELOS DE REDES NEURAIS ARTIFICIAIS 40
4.2.1. PERCEPTRON DE MÚLTIPLAS CAMADAS (MLP) 41
4.2.2. ELM (*Extreme Learning Machine*) 42
4.2.3. ESN (*Echo State Network*) 44
4.2.4. Híbrido ESN-ELM 46
4.3. ALGORITIMOS DE OTIMIZAÇÃO BASEADOS EM ENXAME 47
4.3.1. ABC (*Artificial Bees Colony)* 48
4.3.2. PSO (*Particle Swarm Optimization*) 51
4.4. CONSIDERAÇÕES 52
5. CONJUNTO DE AMOSTRAS E METODOLOGIA 54
5.1. INTRODUÇÃO 54
5.2. AMOSTRAS EXPERIMENTAIS 54
5.3. IDENTIFICAÇÃO DO SISTEMA DINÂMICO DO LASER DE PONTOS QUÂNTICOS 61
5.3.1. Entradas e Sinal de Excitação 61
5.3.2. Arquitetura 62
5.3.3. Representação Matemática 63
5.3.4. Ordem, Estrutura e Complexidade do Modelo NARX 65
5.3.5. Parâmetros do Modelo 65
5.3.6. Avaliação de Desempenho 69
5.4. CONSIDERAÇÕES 70
6. RESULTADOS 71
6.1. INTRODUÇÃO 71
6.2. RESULTADOS PARA DETERMINAÇÃO DE PARÂMETROS 71
6.3. ANÁLISE DOS RESULTADOS GLOBAIS 73
6.3.1. Modelos Generalistas 73
6.3.2. Modelos Especialistas 76
6.3.3. NARX-Especialistas 79

6.4. ANÁLISE DOS RESULTADOS ESPECÍFICOS 86
6.4.1. Análise Gráfica das Predições das Séries Temporais 87
6.4.2. Análise Gráfica do Erro por Séries Temporais 90
6.4.3. Curvas P-I 94
6.5 CONSIDERAÇÕES 95
7. CONCLUSÕES 96
7.1. SUGESTÕES PARA TRABALHOS FUTUROS 99
REFERÊNCIAS 100
SOBRE O AUTOR 108

1 Introdução

1.1.RELEVÂNCIA DO TEMA

Laser de pontos quânticos semicondutores (LPQ) compõem uma recente classe de dispositivos que se apresenta como alternativa para aplicações em telecomunicações e na medicina. Durante o desenvolvimento de uma nova fonte de laser, um passo importante consiste na modelagem dos dispositivos a serem construídos. Isto requer o uso de bons métodos, capazes de incorporar os efeitos causados por vários dos fenômenos físicos presentes em dispositivos reais, permitindo que o modelo desenvolvido auxilie na interpretação de resultados experimentais ou na sua previsão.

As modelagens mais usuais do LPQ são o modelo baseado em descrição de microestados e o modelo baseado em equações de taxa (KANTNER, MITTNENZWEIG e KOPRUCKI, 2017), porém ambos necessitam de um grande esforço computacional devido à complexidade da modelagem. O primeiro método utiliza mecânica quântica na descrição do sistema (GRUNDMANN, HEITZ, *et al.*, 1997) e o segundo, mais simples do que o primeiro, utiliza uma descrição de dinâmica de populações (SUGAWARA, HATORI, *et al.*, 2005).

Neste contexto, com a finalidade de se obter uma predição precisa, comparável as duas modelagens tradicionais, porém com um custo computacional substancialmente menor, o que facilita o desenvolvimento e disseminação da tecnologia. Este trabalho propõe realizar a predição através de uma modelagem com redes neurais artificiais (RNA) para a estimação da não linearidade de um Modelo Não-Linear Autoregressivo com Entradas Exógenas *(Nonlinear AutoRegressive models with eXogenous inputs* — NARX). Foram utilizados modelos como o Perceptron de Múltiplas Camadas (*Multilayer Perceptron* — MLP) (AGUIRRE, 2015), Máquina de Aprendizado Extremo (*Extreme Learning Machine* — ELM) (HUANG, ZHU e SIEW, 2006) e Redes com Estados do Eco (*Echo State Network* — ESN) (JAEGER, 2010), além de modelos híbridos ESN, como os modelos ABC-ESN e PSO-

ESN, que utilizam os algoritmos de otimização Colônia Artificial de Abelhas (*Artificial Bees Colony* — ABC) (KARABOGA, 2005) e Otimização por Enxame de Partículas (*Particle Swarm Optimization* — PSO) (KENNEDY e EBERHART, 1995), e o modelo híbrido ESN-ELM (BUTCHER, VERSTRAETEN, *et al.*, 2010) visando incrementar os resultados obtidos com a ESN.

1.2.OBJETIVO GERAL

O objetivo deste trabalho é propor modelos representativos da dinâmica não linear de lasers de pontos quânticos semicondutores utilizando uma modelagem por redes neurais artificiais.

1.3.OBJETIVOS ESPECÍFICOS

Os objetivos específicos desse trabalho são:

- Estudar o comportamento dinâmico de LPQ para diferentes configurações experimentais;
- Estudar técnicas de modelagem para sistemas não lineares (e.g., NARX);
- Estudar os mais recentes modelos de RNAs e verificar quais tipos mais se adequam para a modelagem do LPQ;
- Desenvolver uma ferramenta de software utilizando RNA para a simulação de lasers de pontos quânticos em modo contínuo que possa ser rapidamente treinada a partir dos dados experimentais e também gerar um resultado de predição mais rapidamente que os modelos tradicionais.

1.4.JUSTIFICATIVA

Tradicionalmente, a comunidade científica trata o problema da modelagem de LPQ a partir de duas principais vertentes a saber: o modelo baseado em descrição de microestados (GRUNDMANN, HEITZ, *et al.*, 1997) e o modelo baseado em equações de taxa (SUGAWARA, HATORI, *et al.*, 2005).

Embora leve a resultados mais precisos em termos de capacidade de predição, o modelo baseado em descrição de microestados envolve uma descrição

que recai na mecânica quântica, sendo, portanto, preferido entre a comunidade de físicos (SANTOS, SIMAS FILHO e THÉ, 2016).

O modelo baseado em equações de taxa, por outro lado, é substancialmente mais simples e intuitiva, pois consiste numa descrição de dinâmica de populações, sendo mais utilizada na comunidade de engenheiros (SANTOS, SIMAS FILHO e THÉ, 2016).

O inconveniente desta segunda abordagem recai justamente na capacidade de reproduzir resultados ou de prever comportamentos. Uma vez que isto é indispensável para, e.g., realizar procedimentos de extração de parâmetros, auxiliar no projeto de dispositivos, ou mesmo em ferramentas de simulação assistida em ambiente CAD. Na prática o que se faz é aprimorar a descrição do sistema físico mediante aumento do número de equações de estado que descrevem o dispositivo (THÉ, 2010). Inevitavelmente, isto leva a um significativo aumento do esforço computacional nas tarefas de estudo numérico de desempenho do laser, o que dificulta sobremaneira a obtenção de resultados confiáveis com pouco tempo de processamento.

Justamente nesse contexto, o presente trabalho encontra espaço, pois busca por modelos alternativos de aprendizado de máquina capazes de reproduzir e/ou prever as características de desempenho de LPQs em operação contínua e pulsada.

Uma alternativa a se buscar neste contexto, é a adoção de modelos que bem representassem a dinâmica não-linear existente na resposta destes dispositivos às entradas em degrau e impulso, de tal modo que seu comportamento temporal pudesse ser previsto. Isto seria importante porque a curva Potência vs. Corrente (P-I) poderia ser simulada, além do que, as características dinâmicas de desempenho (como sobressinal, tempo de acomodação, fator de amortecimento etc.) poderiam ser estimadas, juntamente com as características dos pulsos (largura a meia-altura, potência de pico, taxa de supressão), em caso de investigação de operação pulsada. Em outras palavras, isto significa fazer a identificação do sistema laser a partir de amostras de entrada e saída, as quais são respectivamente entendidas, em primeira aproximação, como a corrente elétrica de alimentação do dispositivo, e a potência ótica resultante do LPQ.

A consideração mencionada implica conceber o laser sob acionamento direto como um sistema de entrada múltipla e saída única (*Multiple Input - Single Output*

— MISO) (AGUIRRE, 2015), que pode ser bem descrito em tempo discreto pelo modelo NARX (AGUIRRE, 2015), um mapeamento não-linear que utiliza como entradas os valores atuais e passados do sinal de entrada e os valores atrasados do sinal de saída, produzindo uma estimativa da saída atual do sistema (ALAZRAI e LEE, 2012).

Uma abordagem para a estimação da não-linearidade do NARX é utilizar uma RNA e, mais especificamente, no presente trabalho, foi investigada utilização de arquiteturas de RNA do tipo ELM, ESN, ABC-ESN, PSO-ESN, ESN-ELM e MLP, perpassando pelo ciclo do fluxograma simplificado (Figura 2.1), até determinar o modelo ideal para a predição do laser.

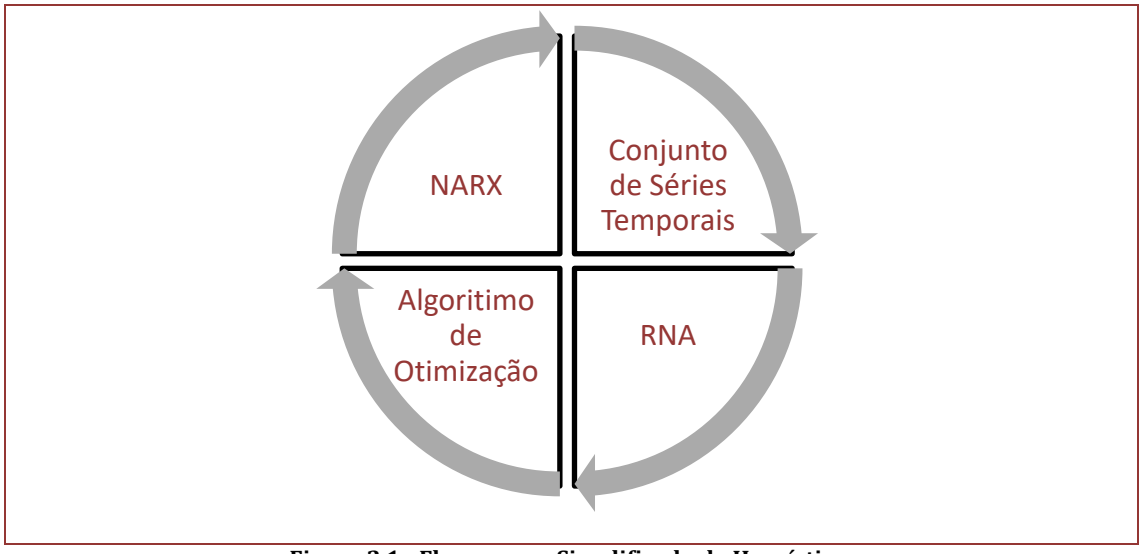


Figura 2.1 - Fluxograma Simplificado da Heurística.

1.5.ORGANIZAÇÃO DA DISSERTAÇÃO

Esta dissertação está organizada em sete capítulos e encontra-se com uma distribuição de acordo com a seguinte estrutura:

- **Capítulo 2**: é iniciada a fundamentação teórica, com a definição do LPQ, suas características, aplicações além de exemplificação de modelagens de sua dinâmica já existentes;
- **Capítulo 0**: é apresentada a definição de séries temporais, a problemática da central do trabalho, a predição, e o passo-a-passo para a identificação de sistemas dinâmicos, fundamental para o detalhamento da metodologia do trabalho;

- **Capítulo 0**: detalha as três diferentes RNAs utilizadas no trabalho: ELM, ESN e MLP, além do modelo híbrido ESN-ELM, e adicionalmente os algoritmos baseados em exames, utilizados no trabalho em uma tentativa de otimizar os resultados obtidos com a ESN;
- **Capítulo 0**: é dedicado à apresentação dos dados experimentais utilizados no trabalho, além da metodologia usada;
- **Capítulo 0**: são discutidos os resultados obtidos com os modelos propostos, iniciando-se pela avaliação dos resultados globais, e posteriormente é realizada uma análise dos resultados específicos do melhor modelo escolhido;
- **Capítulo 0**: são expostas as conclusões da dissertação além de recomendações para trabalhos futuros que poderão apoiar-se nos resultados obtidos e ampliar sua gama de aplicações;
- No Anexo são apresentados os resumos dos trabalhos publicados oriundos da pesquisa realizada.

2 Laser semicondutores de pontos quânticos

2.1.INTRODUÇÃO

O tópico presente inicializa a fundamentação teórica necessária a este trabalho, com a definição do laser semicondutor de pontos quânticos, suas características e aplicações (e.g., geradores de pulsos óticos, lasers de alta potência etc.), assim como, a apresentação de estudos que resultaram nas diversas modelagens existentes

2.2.DEFINIÇÕES E CARACTERÍSTICAS

Os semicondutores distinguem-se de outros materiais por terem na estrutura eletrônica uma faixa de energia (*gap*) entre a faixa de energia ocupada mais alta (a banda de valência) e a banda de energia mais baixa não ocupada (a banda de condução), ilustrada na Figura 2.1. *Bandgap* de um semicondutor é medido entre o topo da banda de valência e o fundo da banda de condução e é chamado direto se esses extremos coincidirem no momento (e.g., GaAs) e chamados de indiretos se não coincidirem (Si) (HOOGLAND, 2008).

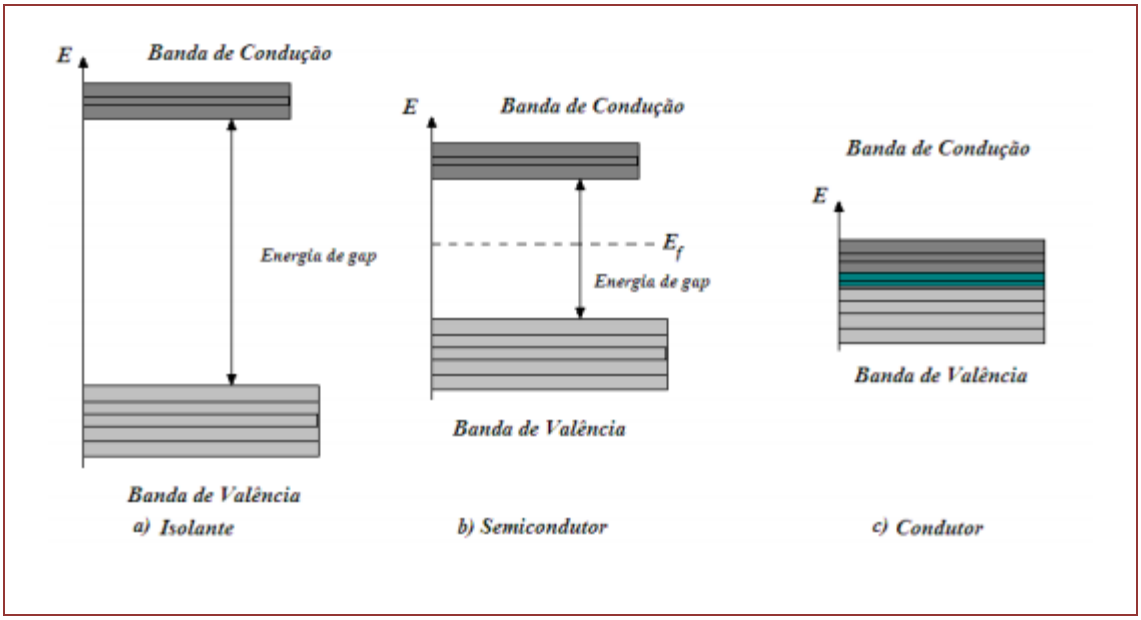


Figura 2.1 - Classificação dos sólidos com relação à banda de energia. Fonte: (SILVA , 2010).

Há uma variedade de diferentes semicondutores e a Figura 2.2 ilustra as características de quatro arranjos estruturais distintos: *bulk*, poços quânticos, fios quânticos e pontos quânticos. Na Figura 2.2 (a), os portadores podem se movimentar livremente nas três dimensões espaciais, não apresentando confinamento quântico, logo sua densidade de estados é continua. Entretanto, na estrutura ilustrada na Figura 2.2 (b) há restrição na movimentação dos portadores em uma das dimensões, resultando na alteração da densidade de estados, fazendo-a deixar de ser contínua. Assim, a terceira estrutura (Figura 2.2 (c)) é o resultado do confinamento em duas dimensões e o ponto quântico (Figura 2.2 (d)) seria a limitação nas três dimensões espaciais (ALVES, 2016).

Assim, o modelo estrutural de semicondutor de ponto quântico (PQ) é definido como qualquer sistema capaz de confinar portadores de carga nas três dimensões espaciais, sendo seu tamanho na ordem de nanômetros, possuindo impacto direto nas suas propriedades ópticas, como, por exemplo, quanto menor seu tamanho, maior o espaçamento entre seus níveis de energia, se tornando semelhante a um átomo (ALVES, 2016).

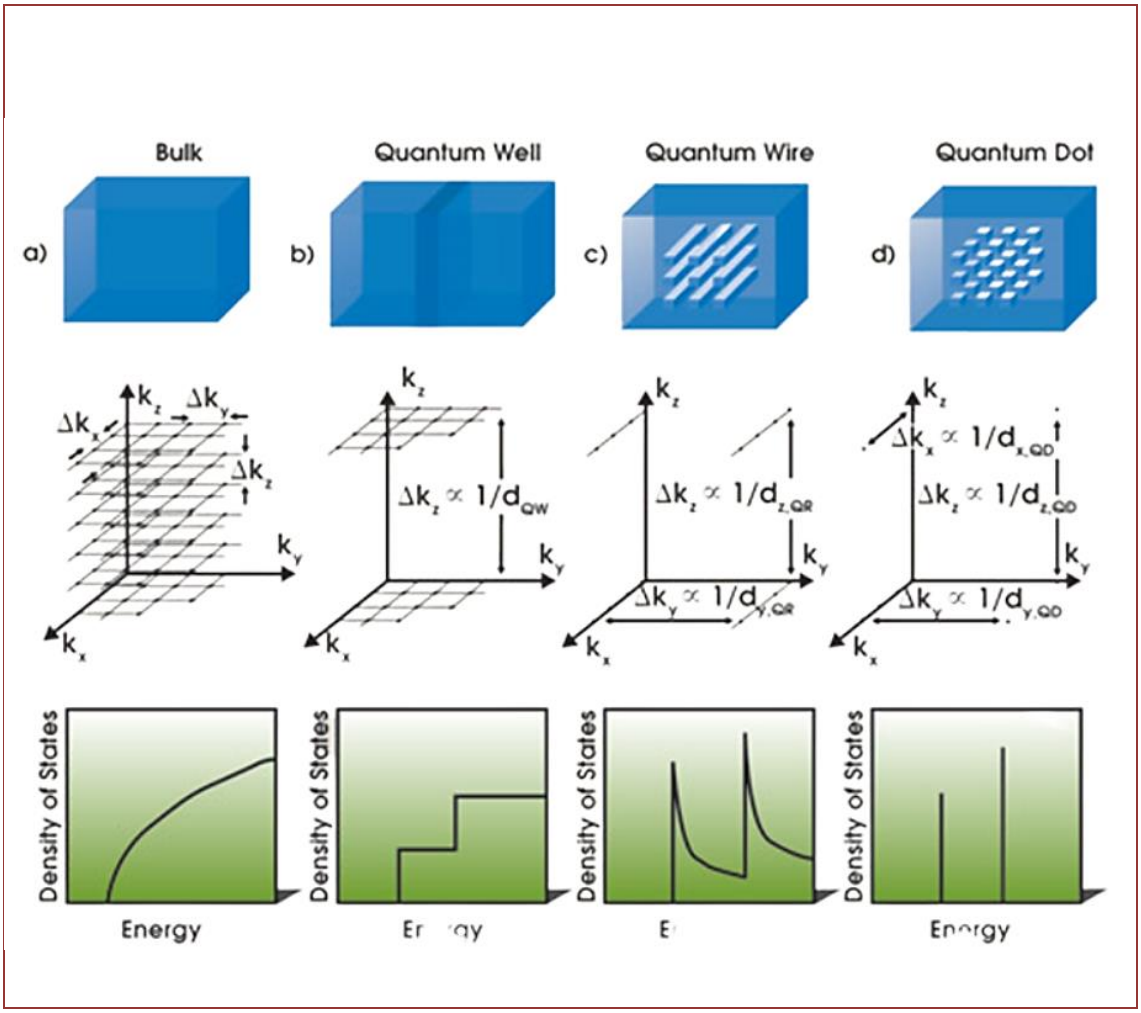


Figura 2.2 - Diferença das características entre materiais semicondutores do tipo *bulk*, poço quântico, fio quântico e ponto quântico. Fonte: (HOOGLAND, 2008).

Um forma de visualizar o processo de formação dos pontos quânticos em semicondutores é ilustrada na Figura 2.3. Inicialmente, a Figura 2.3 (a) apresenta separadamente uma representação das ligas do substrato (círculos pretos) e do filme epitaxial (círculos amarelos) com mesma estrutura cristalina e com Parâmetros de Rede (PR) similares (distância entre os círculos vizinhos), demonstrando um crescimento coerente, onde o material depositado adota naturalmente o mesmo PR que o do substrato, fazendo o filme crescer camada por camada, (PAGNOSSIN, 2008).

Em seguida, a Figura 2.3 (b) ilustra o regime de crescimento bidimensional de Frank-van der Merwe (SILVA, 1999), em que o filme epitaxial cresce, sofrendo assim tensionamento mecânico, continuando com sua estrutura cristalina igual ao do substrato, até atingir uma espessura crítica, ilustrada na Figura 2.3 (c), em que a tensão se torna tão grande que não possa mais ser contida pelo filme, fazendo-o se reorganizar, formando pequenas ilhas tensionadas chamadas de pontos quânticos.

Após o surgimento dos primeiros pontos quânticos, o crescimento é alterado para o regime de Stranski-Krastanov (SILVA, 1999), deixando de ser bidimensional para se tornar tridimensional. Já as camadas epitaxiais recém-depositadas que não são utilizadas para formar os pontos quânticos ficam sob a forma de um filme ultra-fino chamado Camada Molhada (*Wetting Layer* — WL) (C. SILVA, 2008).

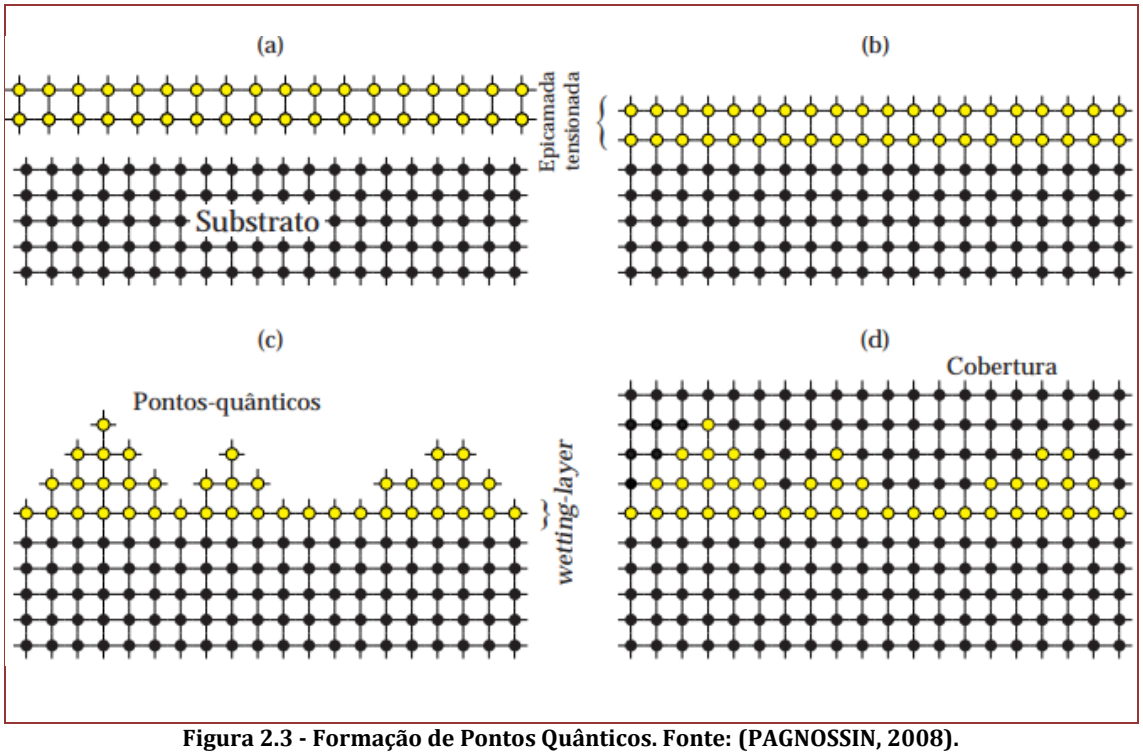


Figura 2.3 - Formação de Pontos Quânticos. Fonte: (PAGNOSSIN, 2008).

Considerando especificamente os lasers semicondutores, dispositivos que possuem a capacidade de emissão estimulada de fótons para criar campos de luz coerentes e de alta energia, em sua forma mais simples, consistem em uma única junção pn dentro de uma cavidade de laser, formada por dois espelhos. Assim que a taxa de geração de fótons excede as perdas ópticas através dos espelhos da cavidade, o campo de luz é amplificado. Devido à diferença do índice de refração entre

materiais semicondutores comuns e ar, atualmente as facetas semicondutoras fornecem refletividade alta o suficiente para permitir o funcionamento do laser, sem a necessidade de cavidades dedicadas com espelhos (LINGNAU, 2015).

Além de fornecer uma fonte confiável de luz monocromática constante e coerente, lasers semicondutores podem exibir uma grande variedade de dinâmicas de portadores eletrônicos diferentes (LARGER e DUDLEY, 2010), (ERNEUX e GLORIEUX, 2010), (LÜDGE, 2012) e (SORIANO, GARCIA-OJALVO, *et al.*, 2013). O modelo mais simples de um laser semicondutor para descrever essas dinâmicas consiste em apenas duas equações (LINGNAU, 2015):

$$\frac{d}{dt}N(t) = J - 2g(N(t) - N_0)S(t) - \frac{1}{T_1}N(t) \tag{2.1}$$

$$\frac{d}{dt}S(t) = 2g(N(t) - N_0)S(t) - 2kS(t) \tag{2.2}$$

O número do transportador de carga $N = N_e + N_h$ descreve o número de elétrons e buracos, respectivamente, na parte ativa do laser e S é o número de fótons. O processo de pulsação é descrito por uma taxa de generalizada J, enquanto as perdas de carga são levadas em conta por um tempo de vida combinado T_1. As perdas óticas são dadas pela taxa de perda de fótons $2k$. O termo $2g(N(t) - N_0)S$ representa a emissão estimulada de fótons, com N_0 sendo o número de portadores de carga necessários para obter transparência do meio. O coeficiente g descreve o ganho linear, proporcional ao coeficiente de Einstein B_k. Assim que o ganho $2g(N(t) - N_0)$ compensa as perdas $2k$, a derivada temporal do número de fótons se torna positiva e a luz do laser é amplificada.

No caso específico do LPQ, a Figura 2.4 ilustra o diagrama esquemático dos níveis de energia e da dinâmica de transporte de cargas, alvo do estudo, onde os portadores são injetados na Heteroestrutura de Confinamento (*Separate Confinement Heterostructure* — SCH), posteriormente migram para a WL, onde relaxam com uma taxa de $\frac{1}{\tau_s}$, em que τ_s é o tempo de difusão no SCH. A partir da WL, que atua como reservatório comum, os portadores podem ser capturados pelos pontos de diferentes tamanhos, a uma taxa de $\frac{1}{\tau_c}$ para o Estado Excitado (*Excited*

State — ES*)*, e em seguida para o Estado Fundamental (*Ground State* — GS) a uma taxa de $\frac{1}{\tau_d}$, sendo τ_c e τ_d os tempos de difusão no ES e no GS, respectivamente.

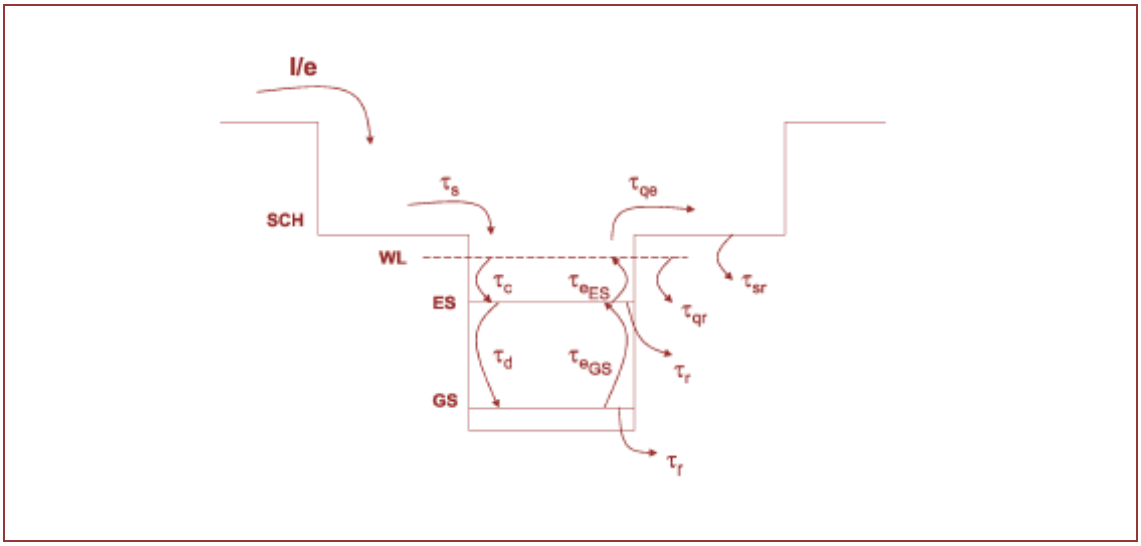


Figura 2.4 – Diagrama esquemático da banda de energia do LPQ na região ativa. Fonte: (GIOANNINI e MONTROSSET, 2007).

Os portadores também podem escapar do GS de volta para o ES, a uma taxa $\frac{1}{\tau_{e_{GS}}}$, e do ES para o WL, a uma taxa de $\frac{1}{\tau_{e_{ES}}}$, em que $\tau_{e_{GS}}$ e $\tau_{e_{ES}}$ são os tempos de saída do GS e do ES, respectivamente. A recombinação pode ocorrer também através de processos radioativos e não radioativos da SCH, WL e com os outros estados de confinamentos, a uma taxa de $\frac{1}{\tau_{sr}}$, $\frac{1}{\tau_{qr}}$ e $\frac{1}{\tau_r}$, sendo τ_{sr}, τ_{qr} e τ_r são os tempos de recombinação no SCH, WL e ES/GS, respectivamente. (GIOANNINI e MONTROSSET, 2007).

Assim, a Figura 2.5 ilustra uma série temporal gerada pela potência ótica de saída dos dois estados, resultante da aplicação de uma determinada corrente de limiar ao LPQ, em degrau após 36,5 ns do início do experimento, influenciada pela temperatura de operação.

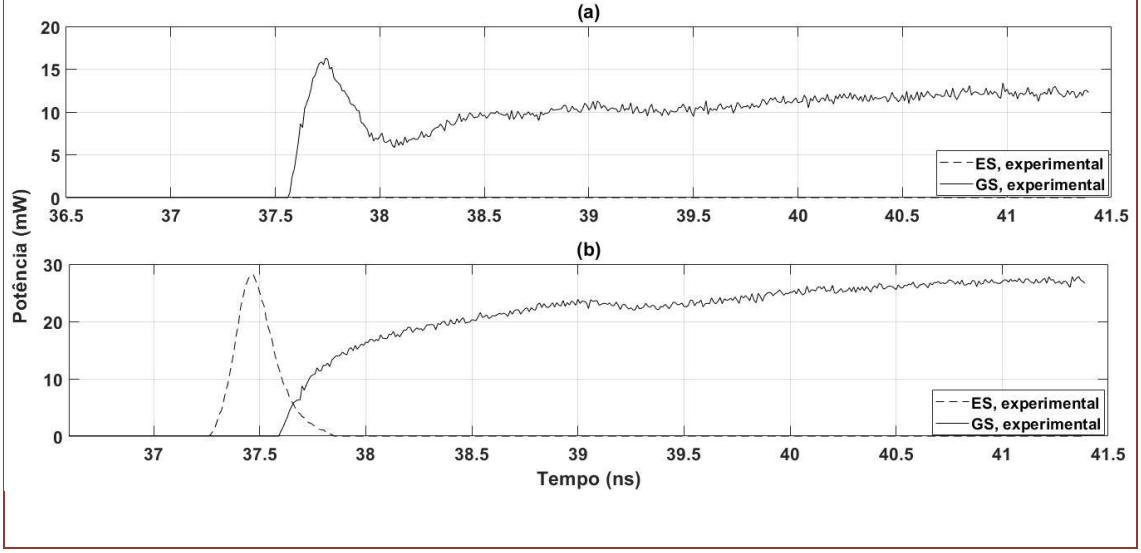


Figura 2.5 - POS a TO = 20ºC. (a) I = 102 mA (b) I = 140 mA. Fonte: Próprio Autor.

2.3.APLICAÇÕES

O fenômeno de quantização do espectro de energias, assim como o confinamento espacial de cargas nas três direções espaciais fazem com que o LPQ ofereça atributos superiores se comparados aos tradicionais dispositivos laser, tais como reduzida corrente de limiar e grande largura de banda (RAFAILOV, CATALUNA e SIBBETT, 2007; BIMBERG, 2003), possibilitando um ampla diversidade de aplicações do LPQ, como por exemplo:

- Moduladores Dinâmicos (SCHWEIZER., JETTER e SCHOLZ, 2003);
- Geradores de pulsos óticos a partir de *q-switching* e *mode-locking* (SUMMERS, REES, *et al.*, 2007; RAFAILOV, CATALUNA e SIBBETT, 2005);
- Lasers de alta potência (KRAKOWSKI, RESNEAU, *et al.*, 2006);
- A possibilidade de emissão a partir dos portadores de ambos os estados confinados permite a obtenção de fontes de banda larga (além de 75 nm) (DJIE, OOI, *et al.*, 2007);
- Fonte óptica para operação atérmica de placas de silício (NISHI, TAKEMASA, *et al.*, 2017);
- Comunicação óptica em datacenter (EZRA, LEMBRIKOV e ZARKOVSKY, 2017).

2.4.MODELAGENS DO LASER DE PONTOS QUÂNTICOS

Para analisar e auxiliar o projeto de fontes laser que tenham pontos quânticos na região ativa, é indispensável a construção de modelos numéricos eficientes, seja do ponto de vista da descrição da riqueza intrínseca dos processos físicos envolvidos, seja do ponto de vista do esforço computacional.

Assim, inúmeros estudos foram desenvolvidos com esta finalidade desde o surgimento desta tecnologia de lasers. A seguir é apresentada uma cronologia de alguns trabalhos importantes nesta área:

- Em 1986 realizou-se uma pesquisa teórica detalhada sobre o ganho e o limiar de LPQs usando a teoria da matriz de densidade (ASADA, MIYAMOTO e SUEMATSU, 1986);
- Em 1996 foi proposto um modelo para simular as características do laser, e equações explícitas para calcular a densidade de limiar (ASRYAN e SURIS, 1996);
- Em 1997 desenvolveu-se o modelo baseado em descrição de microestados utilizando mecânica quântica (GRUNDMANN, HEITZ, *et al.*, 1997);
- Em 2000 foi concebido o modelo de equação de taxa multi-populacional para investigar como o efeito de alargamento homogêneo relacionado à temperatura afeta as características de LPQs (SUGAWARA, MUKAI, *et al.*, 2000);
- Em 2006 apresentou-se um modelo de relaxamento direto da portadora para simular as LPQs Arsenieto de Índio (InAs)/ Fosfeto de Índio (InP) e os resultados simulados foram encontrados em concordância com os experimentais (VESELINOV, GRILLOT, *et al.*, 2006);
- Em 2009 foi proposto o modelo de Circuitos Equivalentes, onde transformaram características do laser em elementos simples de circuitos elétricos, facilitando assim sua modelagem (YAVARI e AHMADI, 2009);
- Em 2017 apresentou-se uma variante do modelo de Circuitos Equivalentes, porém transformando as equações de taxas em elementos de circuitos elétricos (LIN, YUAN, *et al.*, 2017);
- Em 2017 foi desenvolvido o modelo híbrido de equações de taxa utilizando a teoria da termodinâmica (KANTNER, BANDELOW, *et al.*, 2017).

Com este breve resumo de propostas de modelagens da dinâmica LPQ, visualiza-se que até o presente momento o interesse pela tecnologia continua em expansão, havendo a necessidade e espaço para propostas de novos modelos eficientes de predição, foco principal da dissertação.

2.5.CONSIDERAÇÕES

Este capítulo introduziu parte da fundamentação teórica necessária à dissertação, com uma breve apresentação da definição, características e aplicações do LPQ, além de citar o histórico de algumas modelagens propostas desde o desenvolvimento da tecnologia.

3 Modelagem e predição de séries temporais

3.1.INTRODUÇÃO

Neste capítulo serão apresentados aspectos importantes da análise e predição de séries temporais, juntamente com fundamentos de identificação de sistemas.

3.2.SÉRIES TEMPORAIS

Uma série temporal (ST) é qualquer conjunto de observações ordenadas no tempo (MORETTIN e TOLOI, 2006), ou seja, uma série temporal é simplesmente uma sequência de números coletados em intervalos regulares durante um período de tempo. A ordem dos dados é fundamental, diferentemente do que ocorre em modelos de regressão linear, onde a ordem das observações é irrelevante para a análise (MENEZES JÚNIOR, 2012).

Exemplos de STs:

- Valores diários da contaminação das praias de Salvador;
- Valores diários da umidade relativa do ar de Brasília;
- Índices mensais de reajuste do preço da gasolina;
- Precipitação atmosférica anual na cidade de Itaberaba;
- Número médio anual de acidentes de carros nas estradas;
- Variação da maré no Porto da Barra;

As STs podem ser agrupadas nos seguintes tipos:

- **Contínuas:** Possuem valores infinitos entre dois valores quaisquer. O exemplo vi se enquadra nesta categoria;
- **Discretas:** Podendo ser naturalmente discretas, como o valor médio da temperatura diária de uma cidade, ou obtidas através da amostragem de uma ST contínua em intervalos de tempos iguais Δt, essencial para a possibilidade de modelagem e predição da série. Exceto o exemplo vi, todos os demais são deste gênero de ST.

Uma ST discreta pode ser representada pela forma genérica de uma sequência finita de valores de uma certa variável:

$$\boldsymbol{y} \in \mathbb{R}, \{y(1), y(2), \dots, y(N)\} \quad (3.1)$$

sendo N a quantidade de amostras observadas.

E segundo Morettin e Toloi (2006), a ST pode ser decomposta nos seguintes componentes:

1. **Tendência**: mudança do nível médio da série temporal a longo prazo;
2. **Sazonal**: alteração periódica que ocorre em períodos de tempo determinados, como exemplo, fenômenos naturais decorrentes de características meteorológicas;
3. **Cíclica**: oscilação recorrentes, mas sem periodicidade determinada ou regular, deterministicamente previsíveis;
4. **Aleatória**: modificação de natureza imprevisível (ruído).

Consequentemente, segundo Menezes Júnior (2012), as STs podem ser classificadas como:

1. **Determinísticas**: valores futuros podem ser determinados com precisão e absoluta certeza, sem a existência de componente aleatória;
2. **Estocásticas**: Comportamento irregular gerado por um ruído aleatório existente no sistema, definindo-se como uma coleção de variáveis aleatórias ordenadas no tempo.

Assim, ao se analisar uma determinada ST, deve-se focar na necessidade de descrever suas propriedades, como as detalhadas acima, e na determinação das

diversas relações de dependência existentes, guiando-nos para a escolha de um modelo ideal com o qual obtemos as predições de valores futuros com base nos valores passados analisados.

3.3.A QUESTÃO FUNDAMENTAL: PREDIÇÃO

Matematicamente, a predição de uma ST pode ser vista pelo mapeamento $f: \mathbb{R}^p \rightarrow \mathbb{R}$, em que a partir do vetor de regressão (MENEZES JÚNIOR, 2012):

$$\boldsymbol{y}(n) = [y(n), y(n-1), \dots, y(n-p+1)] \tag{3.2}$$

em que $\boldsymbol{y}(n) \in \mathbb{R}^p$, sendo p a dimensão do vetor, deseja-se obter uma estimativa para a saída $y(n+1)$.

A relação do estado atual $\boldsymbol{y}(n)$ e seu próximo valor na ST é dada pela seguinte equação:

$$y(n+1) = f\big(\boldsymbol{y}(n)\big) \tag{3.3}$$

sendo $f(\cdot)$ uma função linear ou não linear.

Assim, a questão pode ser considerada como o problema de encontrar uma função $f(\cdot)$ que fornecerá uma boa aproximação para os futuros valores da ST, através do seguinte modelo:

$$\hat{y}(n+1) = \hat{f}(\boldsymbol{y}(n)) \tag{3.4}$$

sendo $\hat{\boldsymbol{y}}(n+1)$ o valor estimado de $y(n+1)$ e $\hat{f}(\cdot)$ a aproximação de $f(\cdot)$.

Resultando em uma configuração básica de predição:

$$y(n+1) = \hat{f}(\boldsymbol{y}(n)) + e(n) \tag{3.5}$$

sendo $e(n)$ o erro da predição, normalmente utilizado para verificar a qualidade da aproximação, dado por:

$$e(n) = y(n+1) - \hat{y}(n+1) \tag{3.6}$$

A construção de um modelo preditivo para ST perpassa por um ciclo interativo formado pelas seguintes fases (MORETTIN e TOLOI, 2006):

1. **Especificação**: uma classe geral de modelos é considerada para a análise;
2. **Identificação**: podendo ser baseado na análise das funções de autocorrelação e autocorrelação parcial, ou no conhecimento prévio dos dados para definição do modelo mais adequado;
3. **Estimação**: os parâmetros do modelo identificado são estimados;
4. **Validação ou Diagnóstico**: consiste em realizar de testes e verificação dos índices de desempenhos escolhidos.

Portanto, caso o modelo não tenha sucesso, é necessário retornar à fase de identificação, fase mais crítica do processo, segundo Morettin e Toloi (2006), já que é possível que vários modelos distintos se adequem à uma determinada ST, necessitando-se, muitas vezes, utilizar do princípio da parcimônia, ou seja, o uso do modelo mais simples e com menos parâmetros.

3.4.IDENTIFICAÇÃO DE SISTEMAS DINÂMICOS NÃO LINEARES

A modelagem e identificação de sistemas dinâmicos não lineares é uma tarefa desafiadora porque os processos não lineares não compartilham muitas propriedades entre si, sendo assim, o maior problema é a universalidade: a capacidade de descrever uma gama diversa de classes estruturalmente diferentes de sistemas (NELLES, 2013).

Como exemplo simples da tarefa de identificação e sistemas, Nelles (2013) ilustra na Figura 3.1 o processo, assumindo que é um modelo de saída única e múltiplas entradas, sendo $\mu_1(k), \mu_2(k)$ *e* $\mu_p(k)$ os valores das entradas do sistema (e do modelo), no instante k, e $n(k)$ o ruído do modelo, enquanto $y(k)$ a saída do processo e $y'(k)$ a saída do modelo, cujos valores são comparados gerando o erro $e(k)$, que alimenta o modelo para os ajustes dos parâmetros e aumento de sua confiabilidade.

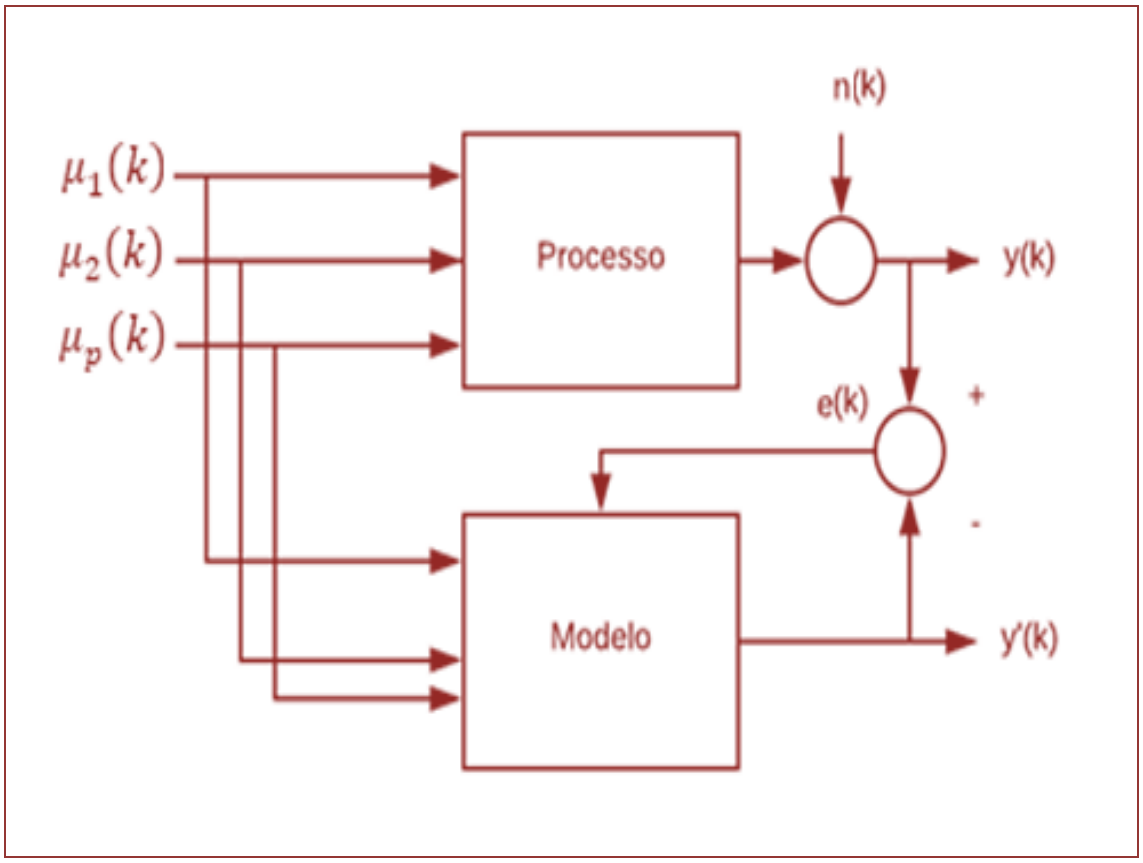


Figura 3.1 - Identificação de Sistemas. Fonte: (NELLES, 2013).

A Figura 3.2 ilustra o fluxograma das principais etapas para o sucesso da identificação do sistema. Inicia-se pela determinação do modelo de entradas, e em seguida é o sinal de excitação do sistema, no caso do laser, é a própria corrente de injeção. Posteriormente a arquitetura do modelo proposto é escolhida, necessitando-se considerar todas as limitações e objetivos do estudo. Sequencialmente, a representação da dinâmica, priorizando os modelos mais simples; a ordem do modelo proposto, evitando o aumento desnecessário de sua complexidade; a estrutura do modelo e sua parametrização. Finaliza-se com sua validação, avaliando-se o desempenho do modelo, analisando se atende todas as expectativas, caso negativo, retorna-se as fases anteriores para novos ajustes e testes.

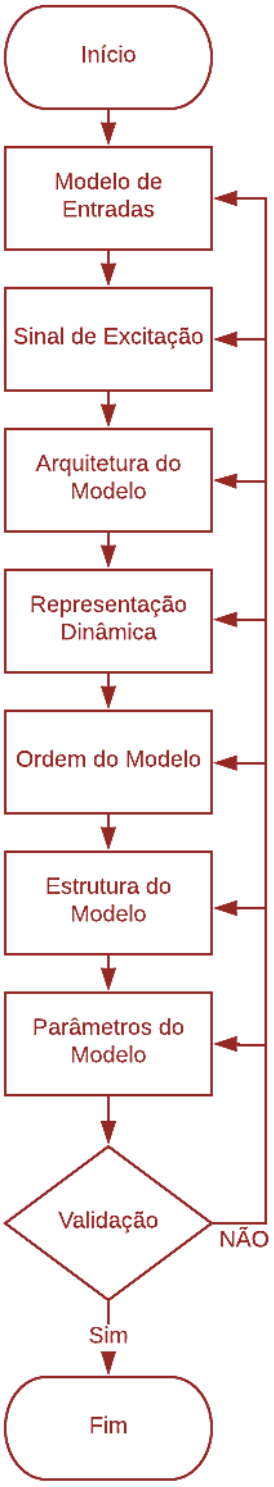


Figura 3.2 - Fluxograma de Identificação de Sistemas Dinâmicos. Fonte: (NELLES, 2013).

Todos as fases são descritas a seguir, detalhando todas as possibilidades consideradas nesta dissertação.

3.5.Modelo de Entrada

Esta primeira fase vai além da mera captação dos dados do sistema a ser estudado, segundo Nelles (2013), é fundamental determinar que informações devem ser priorizadas para sua modelagem e sua formatação, propondo as seguintes possibilidades:

1. **Todas as entradas**: nos casos que sistemas com altas dimensionalidades e grande quantidade de parâmetros envolvidos;
2. **Combinações das entradas**: diminuindo assim a quantidade da estratégia anterior, porém não necessariamente mais prático, em sistemas com muitas variáveis;
3. **Entradas selecionadas sem supervisão**: análise através das componentes principais, descartando entradas de pouca relevância ao sistema;
4. **Entradas selecionadas com supervisão**: análise focada na maximização do resultado, permanecendo apenas variáveis que aumentem a precisão do modelo proposto, porém de complexa realização em processo não lineares.

3.6.Sinal de Excitação

No passo seguinte, o da escolha do sinal de excitação do sistema para posterior coleta das informações de saída, é possível utilizar métodos de identificação determinísticos e estocásticos como o da resposta ao degrau, à função senoidal e ao ruído branco, neste caso é necessário um conhecimento prévio do processo e da proposta de modelo para a escolha mais apropriada (DANTAS, 2013).

É importante armazenar as informações sobre as condições de operação que são cobertas pelo conjunto de dados de treinamento, a fim de poder avaliar a confiabilidade das previsões do modelo e possivelmente evitar, ou pelo menos limitar a extrapolação.

3.7.Arquitetura do Modelo

Na terceira etapa, é necessário definir os critérios, muitas vezes subjetivos e de difícil escolha, para a determinação do modelo apropriado (NELLES, 2013). Abaixo segue uma lista de possíveis critérios:

a) **Aplicação pretendida**: simulação, otimização, controle, detecção de falhas, predição, etc.;

b) **Problema de dimensionalidade**: o número de entradas e saídas relevantes restringe a classe de arquitetura de modelo a ser usada;

c) **Quantidade e qualidade dos dados disponíveis**: dados que são espalhados e ruidosos tendem a ter mais sucesso com modelos globais, que utilizam todo o conjunto de amostras disponíveis. Os modelos globais são mais tendenciosos a média e menos suscetíveis a ruídos, do que locais, que são focados em predizer comportamentos mais específicos, de um determinado subconjunto de dados;

d) **Restrições no tempo de desenvolvimento, treinamento e teste**: tempo de desenvolvimento, que é extremamente dependente do tempo de treinamento, demasiadamente longo pode dificultar o uso prático do modelo, inviabilizando o desenvolvimento da nova tecnologia;

e) **Restrição de Memória**: para muitos processos cuja gama de dados a serem analisados é elevada, é uma restrição de capacidade computacional para a modelagem

f) **Ferramentas disponíveis**: complexidade dos softwares e hardwares compatíveis;

3.8.Representação Matemática da Dinâmica do Sistema

Esta fase do processo de identificação da dinâmica do sistema é caracterizada pela escolha do modelo pretendido, sendo fundamental o prévio conhecimento do sistema para a tomada de decisão (NELLES, 2013).

Assim, Aguirre (2015) descreve os seguintes tipos gerais de representações matemáticas:

a) **Estáticos** (variáveis sem quantificar sua dependência temporal, utilizando-se equações algébricas) e **dinâmicos** (utiliza-se equações diferenciais, cujas variáveis se alteram com o tempo);

b) **Discretos** (representam a evolução do sistema em instantes de tempos discretos e descritos através de equações às diferenças) e **contínuos** (descritos por equações diferenciais e representam a evolução do sistema continuamente no tempo);

c) **Monovariáveis** (representam a relação causa e efeito de apenas um par de variáveis (*Single Input, Single Output* — SISO)) e **multivariáveis** (Única Entrada com Múltiplas Saídas (*Single Input,* M*ultiple Output* — SIMO), Múltiplas Entradas com Única Saída (M*ultiple Input, Single Output* — MISO), Múltiplas Entradas com Múltiplas Saída (M*ultiple Input,* M*ultiple Output* — MIMO));

d) **Determinísticos** (variáveis e parâmetros não são aleatórios) e **estocásticos** (variáveis aleatórias e saída provável dentro de uma função densidade de probabilidade condicional);

e) **Paramétricos** (saídas e parâmetros numéricos) e **não paramétricos** (representações gráficas);

f) **Autônomo** (não contém de forma explicita sinais de entradas gerais) e **não autônomos** (modelo com pelo menos uma variável geral).

Iniciando-se a apresentação de algumas modelagens importantes e comumente utilizadas, segue a equação do modelo linear geral, do qual demais os modelos se originam:

$$y(k) = \frac{B(q)}{F(q)A(q)}u(k) + \frac{C(q)}{D(q)A(q)}v(k) = G(q)u(k) + H(q)v(k) \qquad (3.7)$$

sendo q^{-1} o operador de atraso, de forma que $y(k)q^{-1} = y(k-1)$, $v(k)$ o ruído branco e $A(q)$, $B(q)$, $C(q)$, $D(q)$ e $F(q)$ os polinômios definidos por:

$$A(q) = 1 - a_1 q^{-1} - \cdots - a_{n_y} q^{-n_y} \qquad (3.8)$$

$$B(q) = b_1 q^{-1} + \cdots + b_{n_u} q^{-n_u} \qquad (3.9)$$

$$C(q) = 1 + c_1 q^{-1} + \cdots + c_{n_v} q^{-n_v} \qquad (3.10)$$

$$D(q) = 1 + d_1 q^{-1} + \cdots + d_{n_d} q^{-n_d} \qquad (3.11)$$

$$F(q) = 1 + f_1 q^{-1} + \cdots + f_{n_f} q^{-n_f} \tag{3.12}$$

sendo a_1, a_{n_y}, b_1, b_{n_u}, c_1, c_{n_v}, d_1, d_{n_d}, f_1 e f_{n_f} constantes dos seus respectivos polinômios.

A Figura 3.3 ilustra o modelo geral através de representação esquemática:

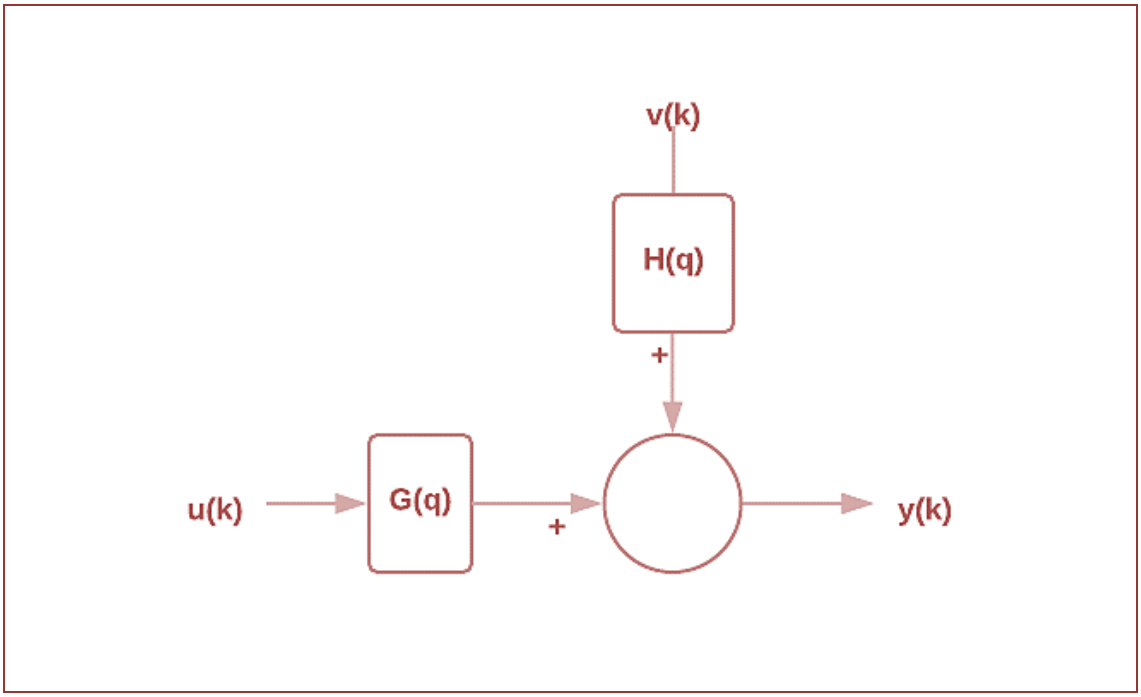


Figura 3.3 - Representação esquemática do modelo geral. Fonte: (AGUIRRE, 2015).

Ao se considerar os polinômios $C(q) = D(q) = F(q) = 1$, na equação do modelo geral (Equação 3.7), obtém-se o modelo Autoregressivo com Entradas Exógenas (*AutoRegresive with eXogenous inputs* — ARX), com a seguinte equação simplificada:

$$y(k) = \frac{B(q)}{A(q)} u(k) + \frac{1}{A(q)} v(k) \tag{3.13}$$

Cuja representação esquemática é ilustrada na Figura 3.4:

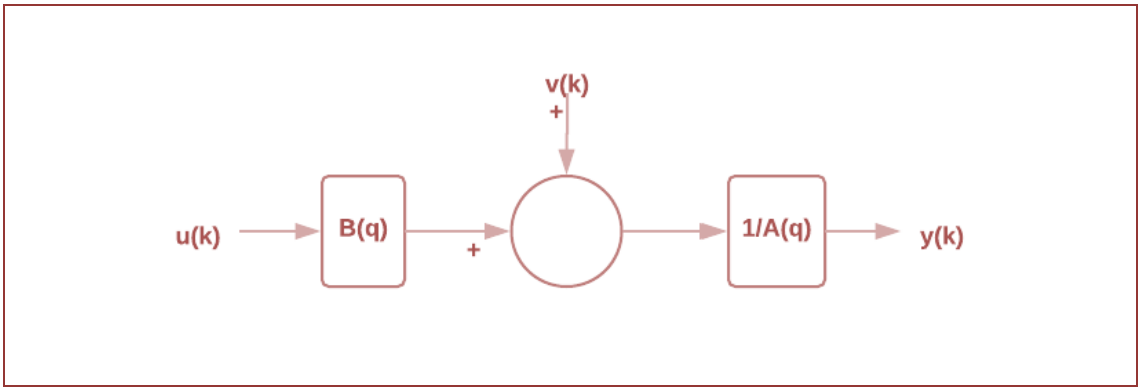


Figura 3.4 - Representação esquemática do modelo ARX. Fonte: (AGUIRRE, 2015).

Já o modelo Autoregressivo (*AutoRegresive* — AR), nada mais é que um caso específico do modelo ARX detalhado acima, porém sem a entrada exógena, resultando na seguinte equação:

$$y(k) = \frac{1}{A(q)} v(k) \tag{3.14}$$

Diferentemente do modelo ARX descrito anteriormente, o modelo Autoregressivo com Média Móvel e Entradas Exógenas (*AutoRegresive Moving Average with eXogenous inputs* — ARMAX), considera apenas os polinômios $D(q) = F(q) = 1$, na equação do modelo geral (Equação 3.7), resultando na equação:

$$y(k) = \frac{B(q)}{A(q)} u(k) + \frac{C(q)}{A(q)} v(k) \tag{3.15}$$

E ilustrada na Figura 3.5:

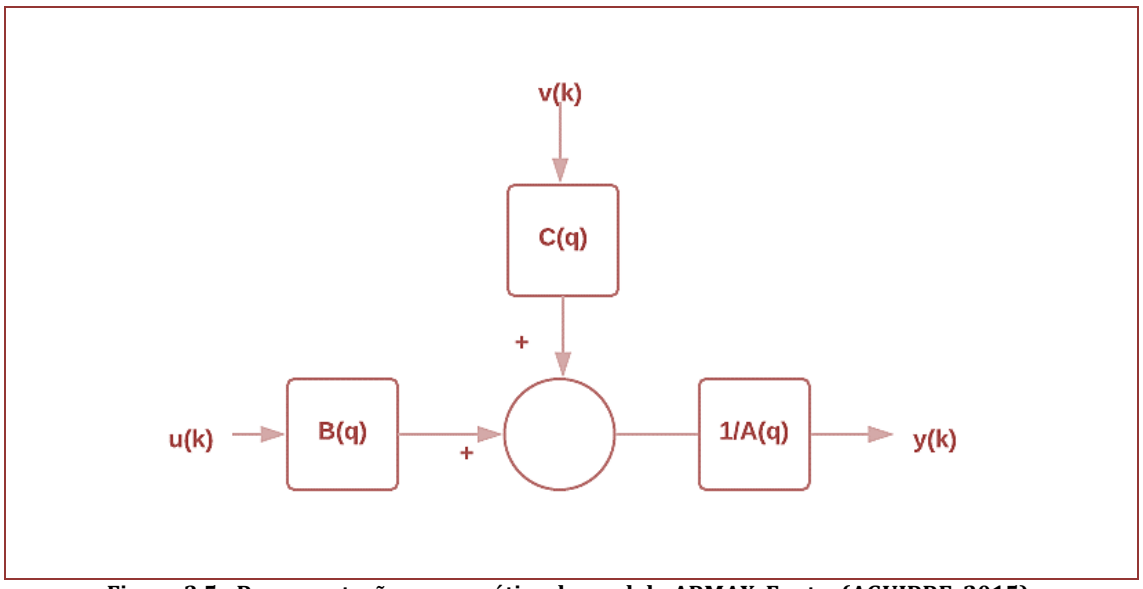


Figura 3.5 - Representação esquemática do modelo ARMAX. Fonte: (AGUIRRE, 2015).

E semelhante ao modelo AR em relação ao ARX, o modelo Autoregressivo com Média Móvel (*AutoRegresive Moving Average* — ARMA) é o caso especifico do ARMAX, porém com a ausência da entrada exógena ($u(k) = 0$), e comumente utilizado em modelagens de STs segundo Aguirre (2015), resultando na seguinte equação:

$$y(k) = \frac{C(q)}{A(q)} v(k) \tag{3.16}$$

3.9.Modelo Autoregressivo Não Linear com Entradas Exógenas

Durante anos, a predição foi dominada por técnicas estatísticas lineares (MENEZES JÚNIOR, 2012), aplicando-se os modelos exemplificados acima, que assumem que a ST estudada é gerada por um processo linear, facilitando a análise e possuindo um baixo custo computacional.

Dependendo da aplicação, a linearização do modelo da ST pode ser suficiente, porém muitas vezes é inadequado devido ao fato gerador da ST ser não linear, ou seja, numa gama de aplicações de STs, modelos lineares não são satisfatórios e as representações não lineares devem ser usadas (ZHANG, PATUWO e HU, 1998).

Para tais casos, um dos modelos possíveis é o modelo Autoregressivo Não Linear com Entradas Exógenas (*Nonlinear AutoRegresive Model with eXogenous inputs* — NARX*)* (ALAZRAI e LEE, 2012) que consiste em um mapeamento não-linear que utiliza como entradas os valores atuais e passados do sinal de entrada (usualmente chamado de entrada exógena) e os valores atrasados do sinal de saída (geralmente chamados de sinais regressivos), produzindo uma estimativa da saída atual do sistema, ou em outras palavras, uma variante do modelo ARX porém com um mapeamento não linear do sistema gerador da ST.

A expressão abaixo, e ilustrada na Figura 3.6, apresenta o NARX para um sistema de entrada única e saída única:

$$y(k) = \phi[y(k-1), \dots, y(k-n); u(k), \dots, u(k-m+1)] \tag{3.17}$$

em que $u(k)$ e $y(k)$ representam os valores de entrada e saída, respectivamente, do sistema no instante k, ao passo que m e n representam a ordem da entrada e da saída, e ϕ representa a função de mapeamento não-linear.

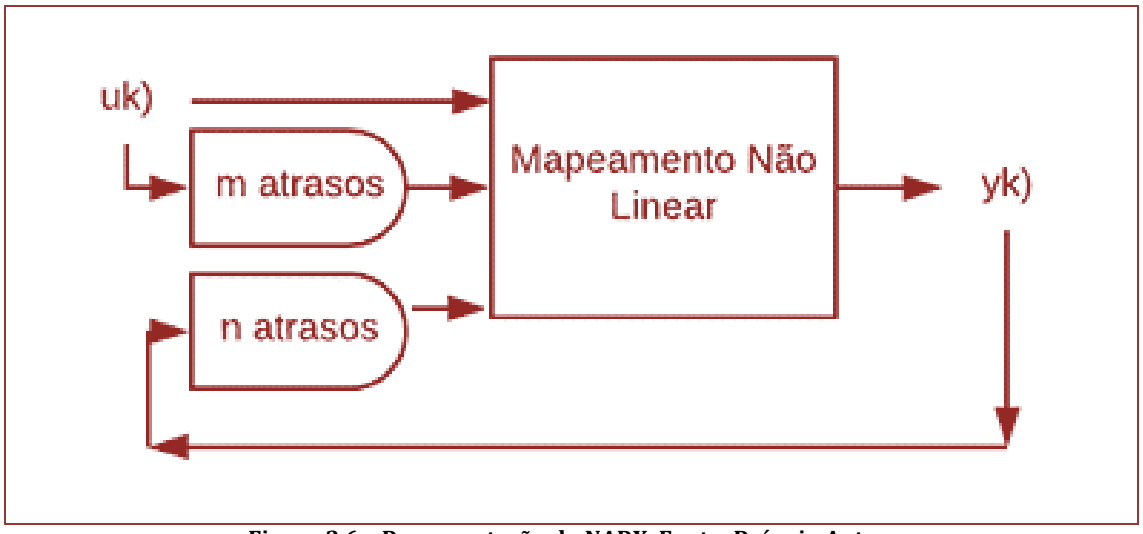


Figura 3.6 – Representação do NARX. Fonte: Próprio Autor.

3.10. Ordem do Modelo

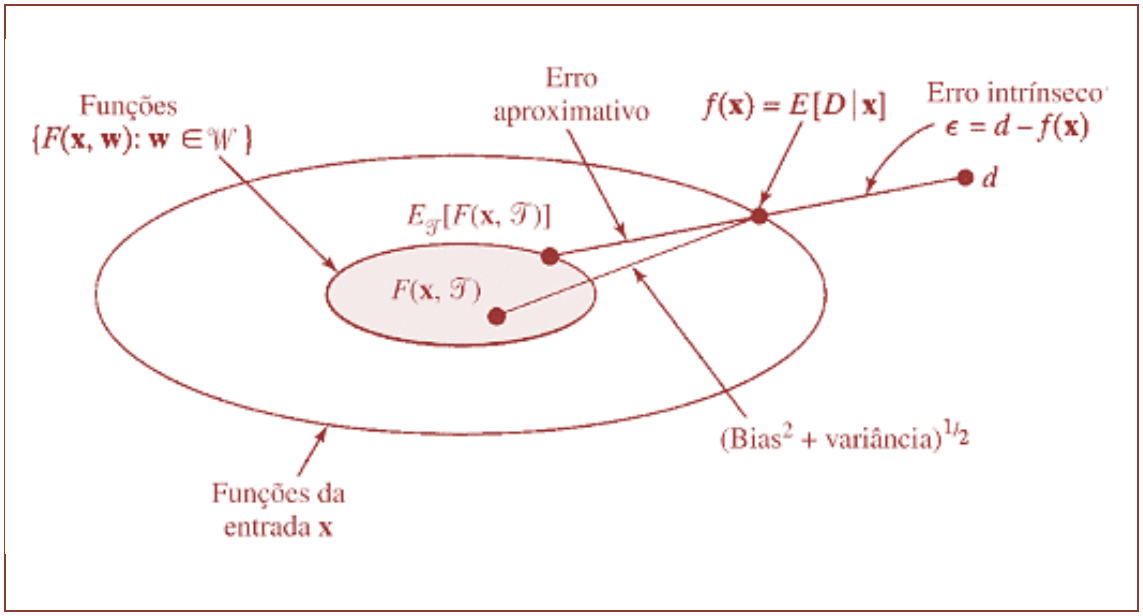
Já na quinta etapa do processo de identificação de sistemas dinâmicos, a escolha da ordem do modelo a ser utilizado é tipicamente realizada através de uma combinação de conhecimento prévio e tentativa e erro. Com a abordagem de dinâmica externa, a escolha por uma ordem dinâmica mais alta do que do modelo aumenta a dimensionalidade do problema e, portanto, sua complexidade.

Comparado com a identificação de sistema linear, ao lidar com sistemas não-lineares, o usuário frequentemente pode ser forçado a aceitar dinâmicas significativamente não modeladas escolhendo uma ordem de modelo muito baixa. A razão para isso é que existe uma compensação entre os erros introduzidos pela dinâmica negligenciada e o erro de aproximação estática. Como a complexidade geral do modelo é limitada pelo dilema de *bias* / variância, uma ordem alta do modelo pode resultar no aumento da imprecisão (NELLES, 2013).

A Figura 3.7 ilustra as relações entre as funções objetivo $f(\cdot)$ e aproximativa $F(\cdot)$, mostrando como os erros estimativos (tendência ou *bias*) e a variância se acumulam. Para a obtenção de um bom desempenho global, é necessário que tanto o *bias* quanto a variância da função aproximativa sejam pequenos. Porém, em rede neurais, somente quando o tamanho da amostra de treinamento se torna infinitamente grande é que pode-se esperar a eliminação da bias e da variância ao mesmo tempo, nos demais casos, o custo para um bias pequeno é uma grande

variância, ou seja, o dilema *bias* / variância, e a consequência é uma convergência excessivamente lente (HAYKIN, 2008).

Figura 3.7 - Ilustração das várias fontes de erro na solução do problema de regressão. Fonte: (HAYKIN, 2008).



Segundo Nelles (2013), modelos de baixa ordem geralmente são suficientes porque o erro do modelo é dominado pelo erro de aproximação causado por uma descrição imprecisa da não linearidade do processo.

3.11. Estrutura e Complexidade do Modelo

Muito mais difícil que a otimização dos parâmetros, a definição da estrutura e complexidade do modelo, pode ser automatizada através de algoritmos evolucionários ou modelos *neuro-fuzzy*, porém, há a dominância da escolha manual devido aos seguintes fatores (NELLES, 2013):

1. Longo tempo de treinamento;
2. Grande número de parâmetros a determinar;
3. Adequação limitada ou inadequação do modelo dinâmico;
4. Limitação de *software.*

3.12. Parâmetros do Modelo

Nesta penúltima etapa, a escolha dos parâmetros do modelo é a mais simples e fácil de se automatizar, com uma grande quantidade de técnicas a serem utilizadas e com várias ferramentas disponíveis. Porém pode ser um grande desafio em casos com um grande número de parâmetros a serem otimizados, também ocorrem casos onde existem várias restrições, múltiplos objetivos e metas o que acaba dificultado o ajuste de parâmetros.

Pode-se utilizar algoritmos de otimização lineares ou não lineares (GOLDBARG e LUNA, 2005) para a tarefa. Enquanto as técnicas lineares são totalmente automatizadas, os algoritmos não lineares, e.g., Colônia Artificial de Abelhas (*Artificial Bees Colony* — ABC) (KARABOGA, 2005) e Otimização por Enxame de Partículas (*Particle Swarm Optimization* — PSO) (KENNEDY e EBERHART, 1995) utilizados no presente trabalho, requerem algumas interações e parâmetros iniciais deverão ser determinados (NELLES, 2013).

3.13. Validação do Modelo

Na última fase, a de validação do modelo, é onde se avalia se todas as decisões tomadas anteriormente obtiveram êxito na modelagem do sistema em questão, sendo um desafio a escolha do critério de avaliação para a tomada de decisão.

A maneira mais fácil de validar o modelo é indicada por Nelles (2013) como a averiguação do modelo nos dados de treinamento, se os resultados já não forem satisfatórios, dificilmente o modelo será aceito, sendo necessário retornar para as etapas anteriores, como sugerido no fluxograma ilustrado na Figura 3.2, até se alcançar o almejado sucesso.

Caso obtenha sucesso, é necessário avaliar posteriormente usando os dados de teste, separados dos de treinamento, de preferência utilizando métricas de desempenho definidas anteriormente, e assim, pode-se ter uma conclusão final sobre o método.

3.5 CONSIDERAÇÕES

Este capítulo mostrou os conceitos básicos de séries temporais, predição e identificação de sistemas dinâmicos, perpassando por todas etapas necessários até o sucesso do modelo escolhido, além da descrição de algumas modelagens lineares e não lineares, especialmente do NARX, utilizado neste trabalho.

Conceitos básicos descritos, porém, importantes, serão utilizados nos próximos capítulos. No capítulo seguinte será apresentada a teoria das redes neurais artificiais e dos algoritmos de otimização usados no trabalho.

4 Introdução às redes neurais artificiais e aos algoritmos de otimização

4.1.INTRODUÇÃO

Este capítulo finaliza a fundamentação teórica necessária, com apresentação de aspectos relacionados à Redes Neurais Artificiais (RNAs), especificando os modelos Perceptron de Múltiplas Camadas (*Multilayer Perceptron* — MLP) (AGUIRRE, 2015), Máquina de Aprendizado Extremo (*Extreme Learning Machine* — ELM) (HUANG, ZHU e SIEW, 2006) e Redes com Estados do Eco (*Echo State Network* — ESN) (JAEGER, 2010). São apresentados também os algoritmos de otimização baseados em enxames, detalhando o Colônia Artificial de Abelhas (*Artificial Bees Colony* — ABC) (KARABOGA, 2005) e a Otimização por Enxame de Partículas (*Particle Swarm Optimization* — PSO) (KENNEDY e EBERHART, 1995).

4.2.MODELOS DE REDES NEURAIS ARTIFICIAIS

Uma abordagem para a estimação da não-linearidade de modelagens matemáticas de sistemas dinâmicos é utilizar uma RNA (AGUIRRE, 2015). RNAs são compostas por camadas de neurônios interconectados, onde a saída de um único neurônio com n entradas, segundo Aguirre (2015), podem ser definidas por:

$$y = f\left(\sum_{j=1}^{n} w_j x_j + b)\right) \tag{4.1}$$

sendo que b (*bias*) e w_j são constantes e $f(\cdot)$ é chamada de função de ativação.

Importante lembrar que em uma RNA multicamada, a saída de um neurônio é a entrada de outro neurônio, da camada subsequente. Sendo assim, a saída de uma rede com um único nó na camada de saída e uma camada oculta é uma função não linear nos parâmetros do tipo:

$$y = f_s \left\{ \sum_{i=1}^{m} w_i f_i \left(\sum_{j=1}^{n} w_{ij} x_j + b_i \right) + b_s \right\} \quad (4.2)$$

sendo $f_s(\cdot)$ a função de ativação do neurônio da camada de saída para uma arquitetura *feedfoward* de única camada oculta (*Single Hidden Layer Feedforward Neural Networks* — SLFNs). Tal função não precisa ser igual a $f_i(\cdot)$, $i = 1, \ldots, m$, que por sua vez, não precisam ser iguais entre si. b_s é o termo de polarização do neurônio da camada de saída, w_i são os pesos da saída de cada neurônio da camada oculta e w_{ij} são os pesos da entrada j, vista pelo i-ésimo neurônio da camada oculta.

4.3.PERCEPTRON DE MÚLTIPLAS CAMADAS (MLP)

A Figura 4.1 mostra o diagrama representativo de uma RNA MLP com duas camadas ocultas e uma camada de saída totalmente conectada, ou seja, todos os neurônios de cada camada da rede estão conectados com todos os neurônios da camada posterior, ressaltando que o sinal segue de maneira progressiva na rede, ou seja, da esquerda para a direita (HAYKIN, 2008).

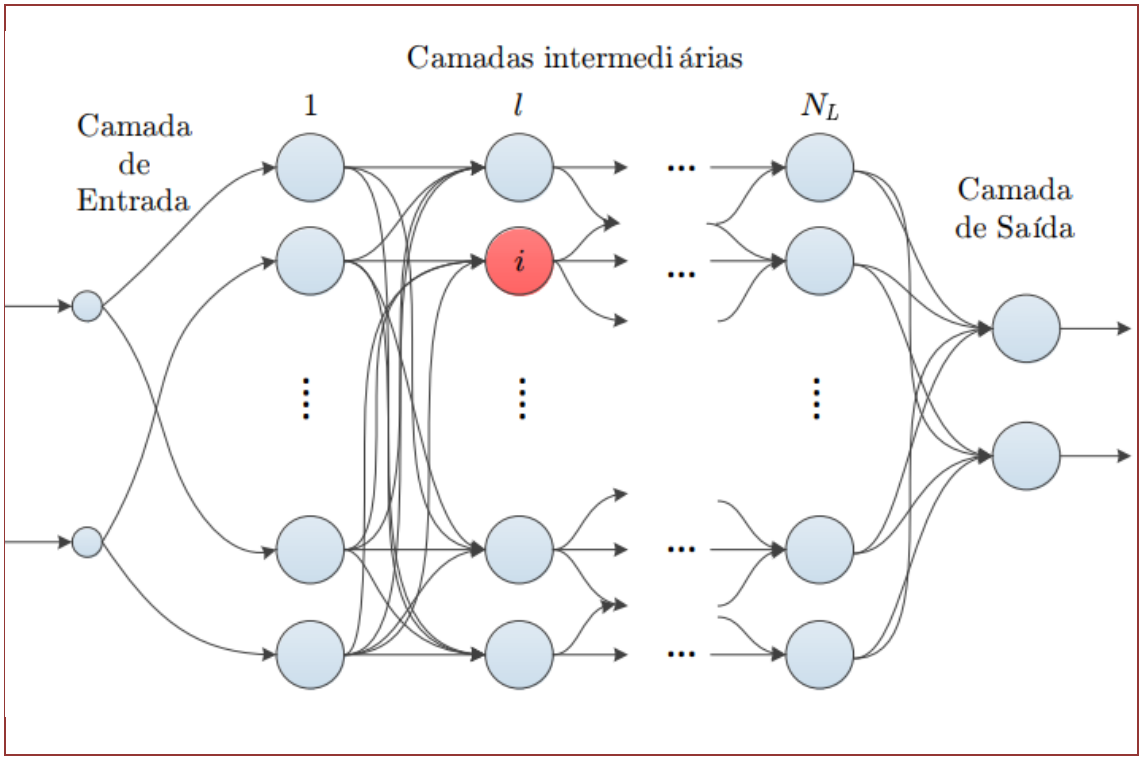


Figura 4.1 - Diagrama de uma RNA MLP. Fonte: (BOCCATO, 2013).

Vale ressaltar que uma rede MLP deve ser treinada para a definição dos pesos sinápticos ideais, e assim, obter o resultado esperado na modelagem. Um método popularmente utilizado é o algoritmo de retropropagação do erro (*error backpropagation*) (HAYKIN, 2008), que possui basicamente duas fases: a primeira onde os pesos sinápticos da rede são fixados e a entrada se propaga através da rede, camada por camada, até a saída. Nesta fase as mudanças são confinadas na ativação potencial e das saídas dos neurônios da rede.

Na segunda fase, ocorre a retropropagação, onde um sinal de erro é produzido e comparado com a saída desejada da rede. O sinal de erro resultante é propagado através da rede, novamente camada por camada, mas desta vez a propagação é realizada de trás para frente, resultando em ajustes sucessivos nos pesos sinápticos da RNA. (HAYKIN, 2008).

4.4.ELM (Extreme Learning Machine)

Outra abordagem de RNA é a arquitetura de rede supervisionada ELM (HUANG, ZHU e SIEW, 2006), que é uma técnica de aprendizado para treinamento de RNA SLFNs e que aumenta a velocidade de aprendizado através da geração aleatória dos pesos de entrada e dos *bias* da camada oculta (PACÍFICO e LUDEMIR, 2012).

O ELM foi proposto como um SLFN "generalizado" com a seguinte função de saída (HUANG e CHEN, 2007):

$$f_L(x) = \sum_{i=1}^{L} \beta_i h_i(x) = \boldsymbol{h}(x)\boldsymbol{\beta} \qquad (4.3)$$

sendo $\boldsymbol{\beta} = [\beta_1,..., \beta_L]^T$ o vetor de peso que liga os nós ocultos e os nós de saída, e $\boldsymbol{h}(x) = [h_1\ (x), \dots, h_L(x)]$ a função não-linear de mapeamento. A real aplicação de $h_i(x)$ pode ser dada por:

$$h_i(x) = G(a_i, b_i, x), \qquad a_i \in R^d, b_i \in R \qquad (4.4)$$

sendo $G(a_i, b_i, x)$ (com os parâmetros de nós escondidos (a, b)) uma função contínua não-linear.

O algoritmo ELM treina a SLFN em duas fases principais: a primeira relacionada ao mapeamento aleatório das características e a segunda sobre a solução linear dos parâmetros. Na primeira etapa, a ELM inicializa aleatoriamente a

camada escondida para mapear os dados de entrada em um espaço de características através de algumas funções de mapeamento não-lineares, podendo ser qualquer função continua não-linear (Gaussiana, Multiquadrática, tangente hiperbólica, etc....) (HUANG, HUANG, *et al.*, 2015).

Os parâmetros dos nós ocultos da ELM são aleatoriamente gerados, independentemente do treinamento de dados iniciais, através de qualquer distribuição continua de probabilidade.

Na segunda etapa, os pesos são relacionados à camada oculta e a camada de saída, indicados por $\boldsymbol{\beta}$, e são estimados pela diminuição do erro de aproximação observando o erro quadrático, como descrito na equação a seguir:

$$\boldsymbol{\beta} \in R^{Lxm} \min ||\boldsymbol{H\beta - T}||^2 \tag{4.5}$$

sendo $\boldsymbol{H}$ a matriz de saída da camada oculta (gerada aleatoriamente) e $\boldsymbol{T}$ a matriz de dados de treinamento e $||\cdot||$ é a norma de Frobenius (GOLUB e VAN LOAN, 1996).

A solução ótima da Equação 4.5 pode ser encontrada da seguinte forma:

$$\boldsymbol{\beta}^* = \boldsymbol{H}^{\dagger}\,\boldsymbol{T} \tag{4.6}$$

sendo $\boldsymbol{H}^{\dagger}$ a matriz inversa generalizada Moore–Penrose de $\boldsymbol{H}$ (BEN-ISRAEL e THOMAS , 2003).

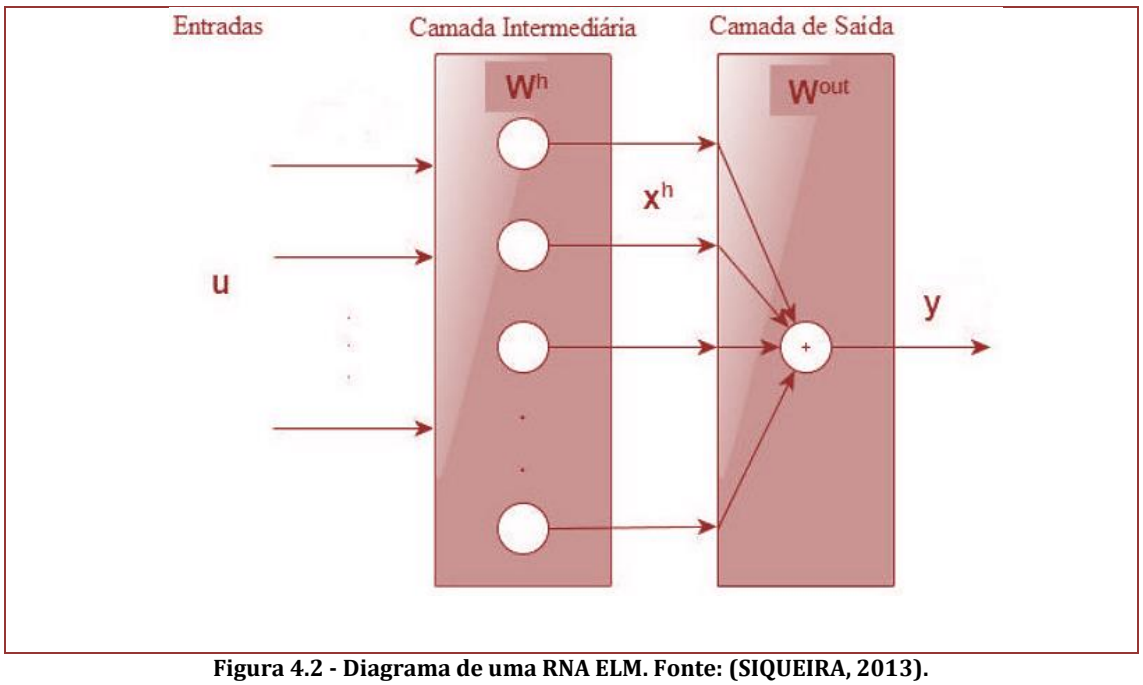


Figura 4.2 - Diagrama de uma RNA ELM. Fonte: (SIQUEIRA, 2013).

Um detalhe importante desta técnica é que apesar da rede ELM oferecer uma boa capacidade de generalização, a escolha aleatória dos pesos sinápticos da camada oculta (pesos de entrada) pode gerar um conjunto não-ótimo de pesos e limiares, ocasionando o efeito de sobreajuste (CAO, LIN e HUANG, 2012).

4.5.ESN (Echo State Network)

A ESN é um novo tipo de RNA recorrente (JAEGER, 2010), que altera apenas os pesos de saída no processo de aprendizado, enquanto os demais pesos permanecem constantes, tornando o treinamento simples, diminuindo sua complexidade e custo computacional, para algumas aplicações é possível obter uma alta eficiência de treinamento quando comparada às demais arquiteturas de RNA, resultando em um amplo uso no processamento de imagens, reconhecimento de fala, previsão de tráfego móvel, entre outros (YANG, ZHU e WANG, 2017).

A ESN sem conexões de realimentação de saída consiste basicamente de três partes resumidas da seguinte forma (QIAO, LI, *et al.*, 2017):

1. K unidades de entradas proporcionam estimulação externa $\boldsymbol{u}(n) \in R^k$;
2. N unidades internas (neurônios), possuindo estados internos $\boldsymbol{x}(n) \in R^N$;
3. L unidades de saída gerando o sinal de saída $\boldsymbol{y}(n) \in R^L$.

Os estados internos $\boldsymbol{x}(n)$ e o sinal de saída $\boldsymbol{y}(n)$ podem ser calculados da seguinte forma:

$$\boldsymbol{x}(n) = \boldsymbol{f}(\boldsymbol{W}_x(n-1) + \boldsymbol{W^{in}}\boldsymbol{u}(n)) \quad (4.7)$$

$$\boldsymbol{y}(n) = \boldsymbol{f^{out}}(\boldsymbol{W^{out}}\boldsymbol{X}(n)) \quad (4.8)$$

sendo $\boldsymbol{X}(n) = [\boldsymbol{x}(n)^T, \boldsymbol{u}(n)^T]^T$ a concatenação dos estados internos e do vetor de entrada, $\boldsymbol{f}(\cdot) = [f_1(\cdot), f_2(\cdot), \dots, f_N(\cdot)]^T$ a função de ativação dos estados internos e $\boldsymbol{f^{out}}(\cdot) = [f_1^{out}(\cdot), f_2^{out}(\cdot), \dots, f_L^{out}(\cdot)]^T$ a função de ativação da saída. $\boldsymbol{W^{in}} \in R^{NxK}$, $\boldsymbol{W} \in R^{NxN}$ e $\boldsymbol{W^{out}} \in R^{L\,x\,(NxK)}$ a matriz dos pesos de entrada, a matriz de pesos do reservatório gerados randomicamente e a matriz de pesos de saída, respectivamente. A última matriz pode ser calculada da seguinte forma:

$$\boldsymbol{W^{out}} = ((\boldsymbol{H}^T\boldsymbol{H})^{-1}\boldsymbol{H}^T\boldsymbol{S})^T \tag{4.9}$$

sendo $\boldsymbol{H} = [\boldsymbol{X}(1), \boldsymbol{X}(2), \dots, \boldsymbol{X}(n_{,max})]^T$ a matriz de estados internos e $\boldsymbol{S} = [\boldsymbol{s}(1), \boldsymbol{s}(2), \dots, \boldsymbol{s}(n_{,max})]^T$ a matriz da saída alvo. A Figura 4.3 ilustra a dinâmica da RNA ESN, descrita nas Equações 4.7, 4.8 e 4.9.

Um importante elemento que não pode ser ignorado em relação a ESN é a necessidade de o reservatório possuir a Propriedade de Estados do Eco (*Echo-State Property* — ESP) os estados do reservatório $\boldsymbol{x}(n)$ devem ser definidos exclusivamente pelo histórico das entradas $[\boldsymbol{x}(n), \boldsymbol{u}(n-1), \dots]$ (JAEGER, 2010). A condição suficiente para ESP é que $\beta_N < 1$, sendo β_N o maior valor singular da matriz do reservatório $\boldsymbol{W}$, e a condição necessária para ESP é que $\lambda \leq 1$, sendo λ o autovalor da matriz $\boldsymbol{W}$ com o maior valor absoluto (YANG, ZHU e WANG, 2017).

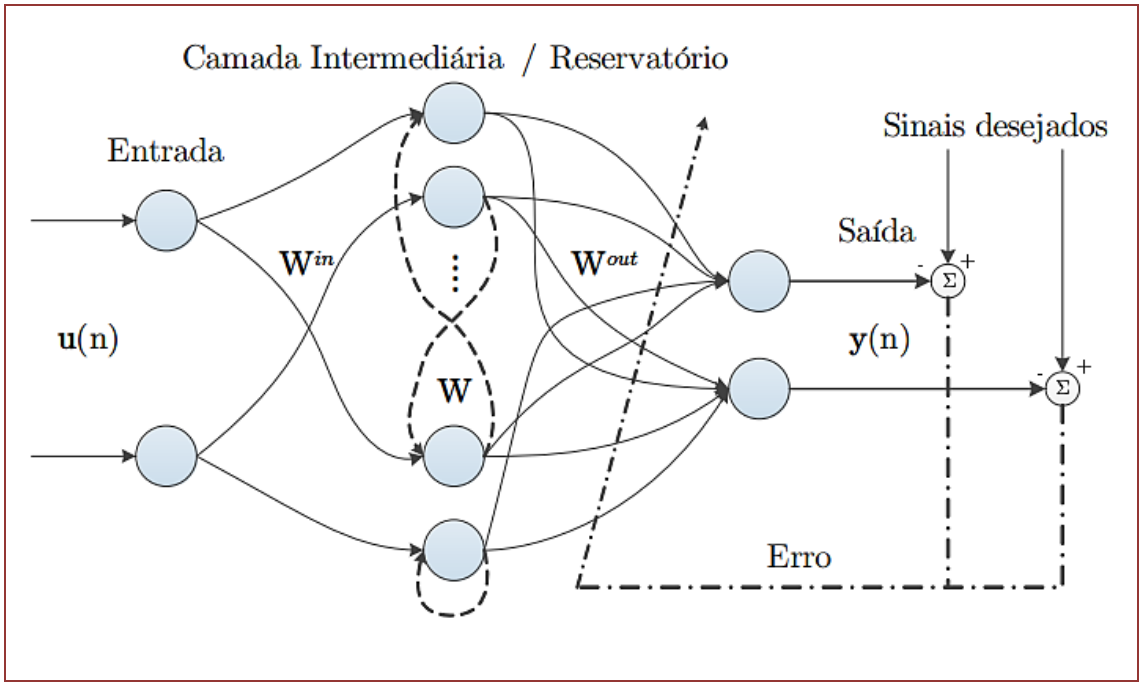


Figura 4.3 - Diagrama de uma RNA ESN. Fonte: (BOCCATO, 2013).

Adicionalmente, para o funcionamento da ESN em suas diversas aplicações, é vital a escolha adequada dos parâmetros de inicialização. A seguir são apresentados os principais parâmetros iniciais da ESN (FAN, WANG e JIN, 2017):

a) Tamanho do reservatório: o número de neurônios internos dentro o reservatório da ESN, sendo provavelmente o aspecto mais determinante, porque o tamanho (N) do reservatório impacta diretamente no número máximo de conexões sinápticas (N2) na rede do reservatório. Reservatórios muito pequenos resultarão na incapacidade de modelar com precisão, enquanto os muito grandes podem gerar problemas com o ajuste excessivo aos dados;

b) Conectividade (σ)**:** o número fracionário de conexões recorrentes dentro do reservatório afeta diretamente a complexidade do reservatório. Muitas conexões podem afetar a capacidade de treinamento da ESN, porque uma conectividade excessiva pode reduzir a diversidade dos estados no reservatório. Na prática é comum se utilizar valores no intervalo entre [0,01;0,2] (BUSING, SCHRAUWEN e LEGENSTEIN, 2010);

c) Raio espectral (ρ): o comprimento da "memória" do reservatório. Quanto maior o raio espectral, mais as entradas anteriores influenciam na saída atual, assim, para garantir a estabilidade da dinâmica, o tamanho do raio espectral ρ, que é o maior autovalor absoluto da matriz de peso de conexão $\boldsymbol{W}$, tem que ser corretamente controlado.

4.6.Híbrido ESN-ELM

A RNA ESN utiliza um combinador linear (JAEGER, 2010), resultando na vantagem de simplificar a estrutura e obter uma solução ótima analiticamente (SIQUEIRA, 2013). Porém, é possível melhorar a eficiência do resultado final através da utilização de camadas de saída não-lineares, como a utilização de um MLP como camada de saída da ESN, essa arquitetura é chamada RNA híbrida ESN-MLP (MAASS, NATSCHLÄGER e MARKRAM, 2002).

Apesar da maior capacidade de mapeamento ao se utilizar uma RNA híbrida ESN-MLP, melhorando assim seu desempenho, esta proposta faz com que a ESN perca uma de suas principais vantagens, que é exatamente a simplicidade no processo de treinamento, impactando em um maior tempo de processamento para o treinamento da RNA (SIQUEIRA, 2013).

Assim, com o objetivo de melhorar o desempenho da RNA ESN, sem a perda das suas principais vantagens, surgiu o modelo híbrido de RNA utilizando uma ELM como camada de saída, criando um modelo com treinamento supervisionado simples, e baseado apenas no combinador linear do ELM, como ilustrado na Figura 4.4 (BUTCHER, VERSTRAETEN, *et al.*, 2010).

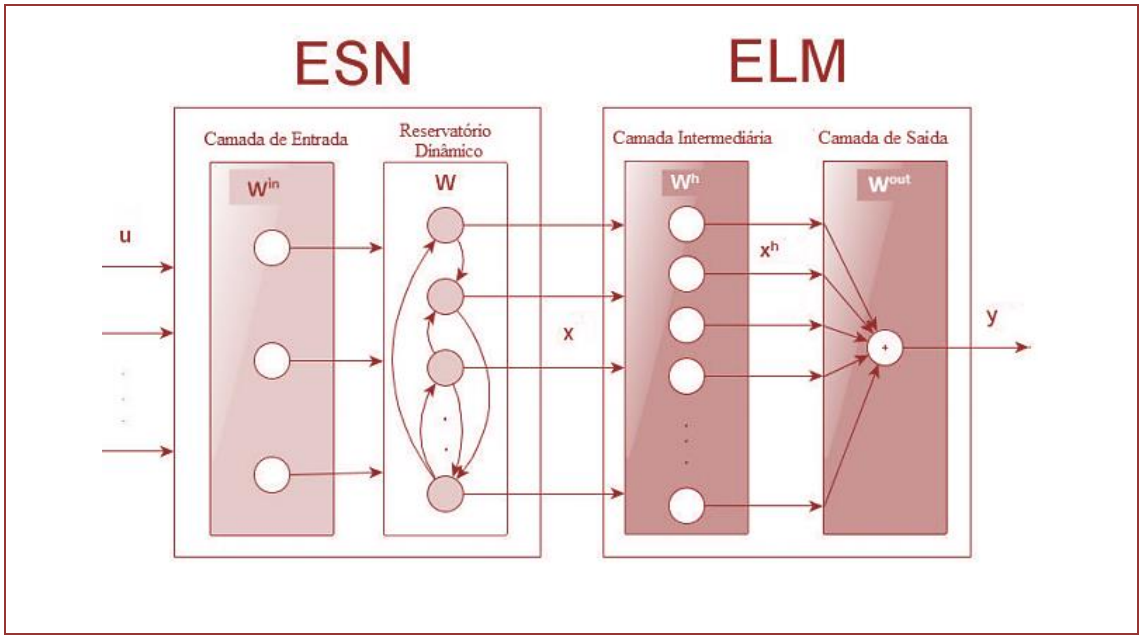


Figura 4.4 - Estrutura Híbrida ESN-ELM. Fonte: (SIQUEIRA, 2013).

4.7.ALGORITIMOS DE OTIMIZAÇÃO BASEADOS EM ENXAME

Com o objetivo de melhorar os resultados obtidos com a RNA ESN, foram utilizados algoritmos de otimização baseados em enxame (*Swarm-based Optimization Algorithms* — SOAs), que utilizam modelos de comportamentos encontrados na natureza para alcançar a soluções ótimas, para determinar os parâmetros iniciais ótimos da ESN, e.g., a conectividade (σ) e o raio espectral (ρ).

A principal diferença entre as SOAs e outros algoritmos, como, o algoritmo de gradiente (*Hill Climbing)* e o Caminhada Aleatória (*Random Walk*), é a possibilidade de utilização de uma população de possíveis soluções para cada iteração ao invés de apenas uma única solução, como os outros algoritmos. Adicionalmente, pode-se empregar diferentes métodos de avaliação para decidir qual solução ótima será adotada, ou seja, escolhe-se apenas a solução que for aumentar o valor encontrado até o momento para a função objetivo (AUGUSTUS, 2009).

Neste trabalho utilizamos dois tipos diferentes de SOAs: Colônia Artificial de Abelhas (*Artificial Bees Colony* — ABC) e a Otimização por Enxame de Partículas (*Particle Swarm Optimization* — PSO).

4.8.ABC (Artificial Bees Colony)

O Algoritmo ABC é um SOA inspirado no comportamento natural da procura de comida por abelhas (KARABOGA, 2005), que possui a vantagem de ser simples e de fácil implementação, sendo amplamente utilizado em estudos na área de comunicação, processamento de imagens e *data mining*, porém possui a característica de cair em uma solução ótima local, podendo apresentar velocidade de convergência e precisão baixas (SHI, ZHAO e SU, 2017).

Na implementação do algoritmo, a colônia é composta pelas abelhas operárias (*employed bees)*, abelhas seguidoras (*onlookers*) e pelas abelhas campeiras (*scouts)*, existindo uma fonte de mel para cada abelha e a quantidade de néctar representa a qualidade ou aptidão em encontrar a solução, como ilustrado no fluxograma da Figura 4.5 e com seus principais passos descritos da seguinte maneira (SHI, ZHAO e SU, 2017):

1. As fontes iniciais da comida são geradas aleatoriamente de acordo com a seguinte fórmula:

$$x_{i,j} = L_j + r_{i,j}(U_j - L_j) \tag{4.10}$$

sendo $x_{i,j}$ a posição da fonte de comida e $i = 1,2, \dots, SN$, SN o número de abelhas operárias; L_j e U_j os limites inferiores e superiores do espaço amostral respectivamente; $r_{i,j}$ um número aleatório no intervalo [0,1] e j um número inteiro aleatório no intervalo [1,D]. D é

a dimensão do problema a ser solucionado. A aptidão é calculada pela fórmula:

$$fit(x_i) = \begin{cases} \dfrac{1}{1+f(x_i)}, f(x_i) \geq 0 \\ 1+|f(x_i)|, f(x_i) < 0 \end{cases} \quad (4.11)$$

sendo $f(x_i)$ a função objetivo da fonte de comida x_i;

2. Abelhas operárias e seguidoras procuram a fonte de comida pela seguinte fórmula:

$$v_{i,j} = x_{i,j} + r_{i,j}(x_{i,j} - x_{k,j}) \quad (4.12)$$

em que k é o número inteiro aleatório no intervalo [1,SN] e $k \neq i$. $r_{i,j}$ é um número aleatório no intervalo [-1,1]. Então calcula-se o valor da aptidão e escolhe-se a aptidão de maior valor da fonte de comida de acordo com o método guloso;

3. As abelhas seguidoras calculam a probabilidade de seleção para fazer a escolha usando:

$$P_i = \frac{fit(x_i)}{\sum_{i=1}^{SN} fit(x_i)} \quad (4.13)$$

se P_i for menor que o número aleatório no intervalo [0,1], a seguidora escolhe a fonte x_i. Procura-se por uma nova fonte de comida em torno de x_i. Seguindo a Equação 4.13 e atualiza-se a fonte de comida de acordo com o método guloso;

4. Após as buscas dos limites, se uma fonte de mel não tiver sido atualizada, a operária desiste da fonte de mel, muda para uma campeira e procura uma nova fonte de mel.

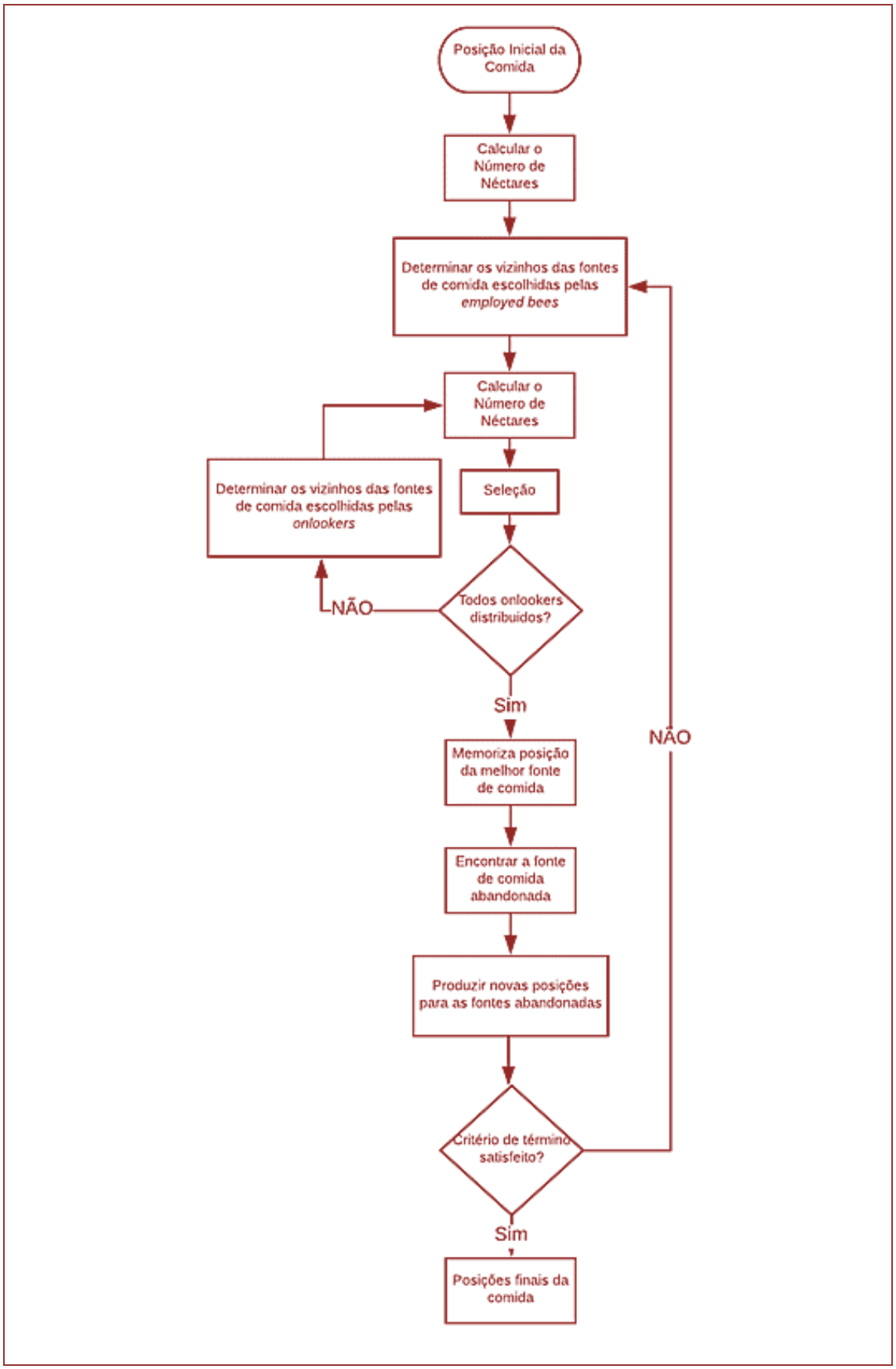


Figura 4.5 - Fluxograma do algoritmo ABC. Adaptado: (LUTHRA, KRISHNA, *et al.*, 2017).

4.9.PSO (Particle Swarm Optimization)

O algoritmo estocástico de otimização PSO foi desenvolvido por Kennedy e Eberhart em 1995 visando resolver problemas complexos de otimização com variáveis contínuas. Inspirado no comportamento social de animais que vivem em enxames, como uma revoada de pássaros ou cardumes de peixes, onde cada indivíduo usa, além de sua própria memória, informações locais fornecidas por seus vizinhos para decidir onde e como se mover (KENNEDY e EBERHART, 1995).

Em outras palavras, o PSO é uma busca estocástica em um problema no espaço n-dimensional, visando minimizar (ou maximizar) a função objetivo em questão. O algoritmo combina métodos de busca local e global (PACÍFICO e LUDEMIR, 2012), possuindo muitas características em comum com técnicas evolucionárias como algoritmos genéticos, porém com as vantagens de ter fácil implementação e poucos parâmetros a serem determinados (AUGUSTUS, 2009).

Segundo Salah (2017), no PSO cada solução única é uma “partícula" que possui valores que são avaliados pela função de objetivo a ser otimizada, cada partícula se movimenta com uma velocidade específica na direção das partículas ótimas atuais em cada interação.

O Fluxograma do algoritmo é apresentado no final deste capítulo na Figura 4.6, e suas principais etapas detalhadas a seguir (PACÍFICO e LUDEMIR, 2012):

1. As partículas são inicializadas aleatoriamente, e mantendo salvos durante o algoritmo valores de suas posições, velocidades e melhores posições alcançadas;
2. Em cada interação, a partícula atualiza a velocidade de acordo com a melhor posição individual ($z_i(t)$) e global ($\hat{z}(t)$), até o momento, de acordo a equação à seguir:

$$v_{ij} = pv_{ij} + c_1 r_1 \left(z_{ij} - d_{ij}\right) + c_2 r_2 (\hat{z}_j - d_{ij}) \quad (4.14)$$

$$d_{ij}(t+1) = d_{ij}(t) + v_{ij}(t+1) \quad (4.15)$$

$$1 \leq i \leq s, 1 \leq j \leq n \quad (4.16)$$

sendo p o peso inercial escalar, r_1e r_2 variáveis aleatórias de uma distribuição uniforme de intervalo [0,1], e os valores $0 \leq c_1, c_2 \leq 2$ coeficientes de aceleração individual e global, respectivamente, parâmetros iniciais do algoritmo;

3. A melhor posição individual visitada até o momento $z(t)$ é atualizada de acordo com a Equação 4.17, enquanto a melhor posição global visitada até o momento $\hat{z}(t)$ é atualizada pela Equação 4.18:

$$z_i(t+1) = \begin{cases} d_i(t+1), \text{se } f\big(d_i(t+1)\big) < f\big(z_i(t)\big), \\ z_i(t), \text{senão} \end{cases} \quad (4.17)$$

$$\hat{z}(t+1) = \underset{z_i(t+1)}{\operatorname{argmin}} f(z_i(t+1)) \quad (4.18)$$

$$1 \leq i \leq s \quad (4.19)$$

A PSO vem sendo aplicada com sucesso em muitas áreas como otimização de funções, treinamento de redes neurais artificiais, como no trabalho de Pacífico e Ludermir (2012), controle de sistemas de lógica *fuzzy*, e outras áreas onde se pode aplicar algoritmos genéticos (AUGUSTUS, 2009).

4.10. CONSIDERAÇÕES

Este capítulo descreveu a teoria das redes neurais artificiais, enfatizando os modelos MLP, ELM e ESN. Foram apresentados também os algoritmos de otimização baseados em enxames, destacando o método fundamentado em colônias de abelhas e o de enxames de partículas, ambos utilizados no presente trabalho.

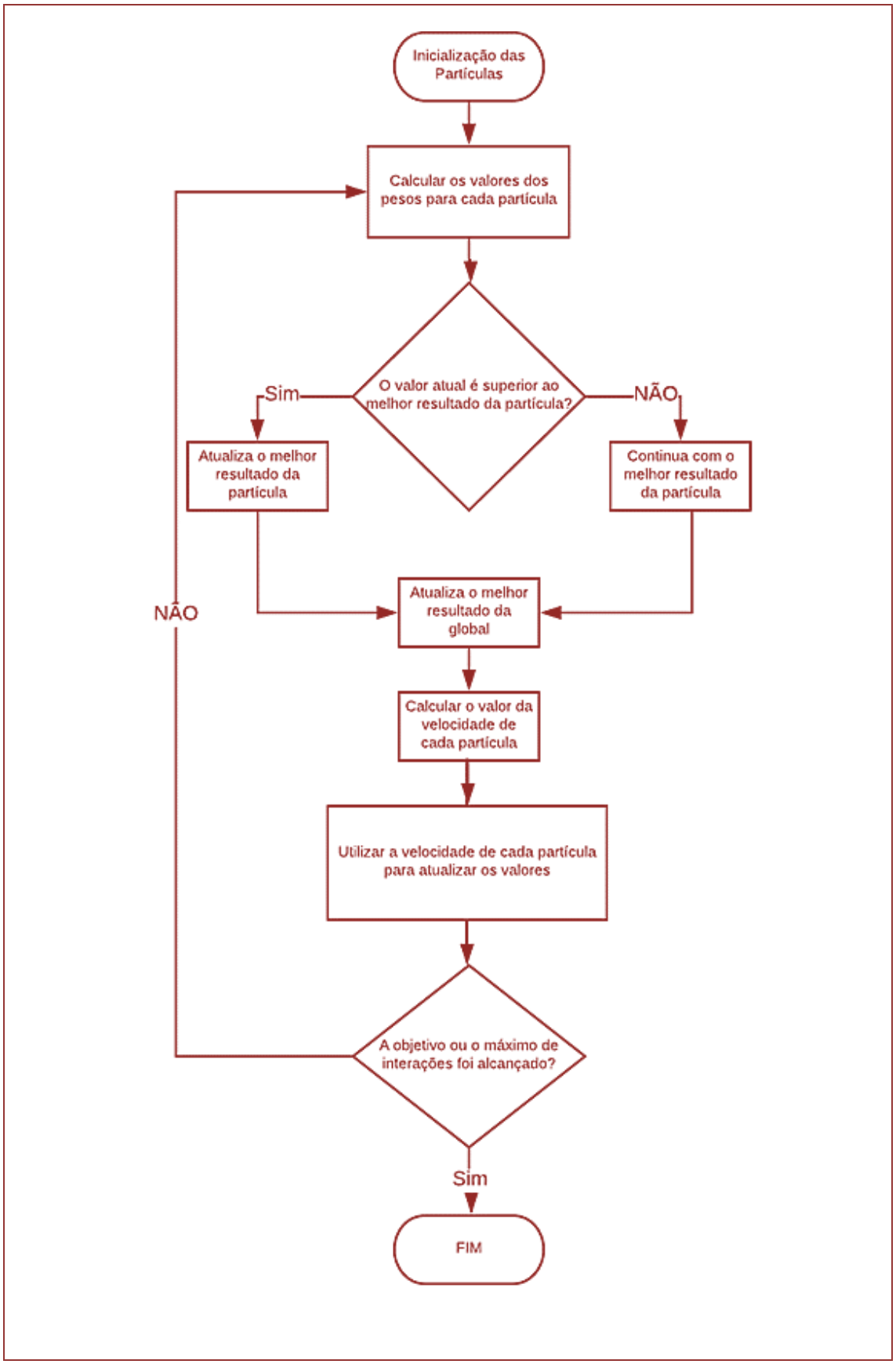


Figura 4.6 - Fluxograma do algoritmo PSO. Fonte: (SALAH, 2017).

.

5 Conjunto de amostras e metodologia

5.1.INTRODUÇÃO

Este capítulo é dedicado à apresentação das amostras experimentais e à definição da metodologia utilizada no trabalho. São apresentadas também as métricas de avaliação empregadas para estimar o desempenho dos modelos propostos, juntamente com a heurística aplicada no processo de escolha da melhor técnica de predição.

5.2.AMOSTRAS EXPERIMENTAIS

Para este trabalho, foram utilizadas as medições realizadas pelo *Semiconductor Optics Group da Technical University of Darmstadt* (Alemanha) com a configuração experimental ilustrada na Figura 5.1, onde no lado esquerdo é ilustrado o laser, em conjunto com o controlador de temperatura e o isolador óptico, enquanto ao lado direito filtro de interferência, o medidor óptico de potência e o fotodetector. Foi medida a potência óptica de saída (POS) filtrada no entorno do comprimento de onda da emissão dos Estado Fundamental (*Ground State* — GS) e Estado Excitado (*Excited State* — ES), para Temperaturas de Operação (TO) iguais à 20°C, 40°C e 50°C, em resposta às excitações do tipo injeção em degrau de corrente.

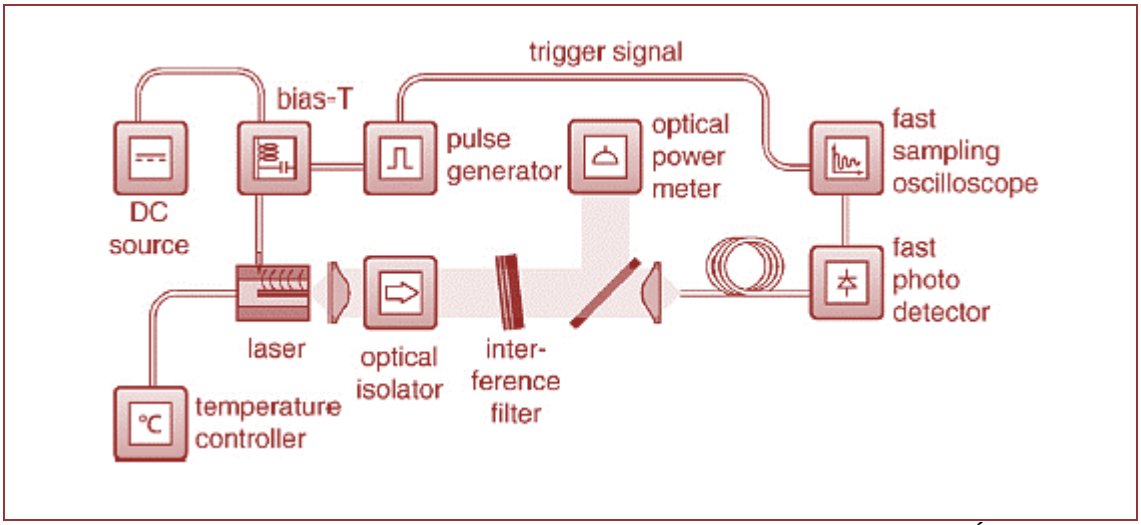


Figura 5.1 - Configuração experimental para medição da POS. Fonte: (DRZEWIETZKI, THÉ, *et al.*, 2010).

A resposta a cada excitação constitui, do ponto de vista da modelagem, uma série temporal que, no experimento de aquisição realizado, possui 510 amostras por curva de Potência Óptica de Saída (POS) resultante de cada corrente de injeção. As amostras são adquiridas entre 36 ns e 42 ns após o início do experimento, com um breve atraso em relação a corrente de limiar, que se inicia em torno de 36 ns. Foram adquiridas um total de 324 séries temporais para diferentes amplitudes do degrau de corrente (aumentando o valor final do degrau de a corrente I_{top} em 1 mA à cada novo experimento), como descrito na Tabela 5.1.

Tabela 5.1 - Características das Amostras Experimentais.

Estado	Temp. (ºC)	$I_{bottom}(mA)$	$I_{top}(mA)$	Quantidades de Séries Temporais
ES	20	60	[64.9; 139.9]	76
GS	20	60	[64.9; 139.9]	76
ES	40	110	[114.9; 185]	71
GS	40	110	[114.9; 185]	71
ES	50	170	[170.9;200]	30

As Figuras 5.2 até 5.8 ilustram a potência ótica de saída dos experimentos para cada par Estado/TO, detalhando a variação da saída ao longo do tempo. Através dessas figuras também é possível constatar que as ST das amostras experimentais

não possuem as componentes de sazonalidade e cíclica, por se tratarem de ST de curta duração, apresentando apenas as componentes de tendência e a aleatória.

Na Figura 5.2, visualiza-se o comportamento da saída do laser no ES à 20ºC, que só consegue ter uma resposta aproximadamente impulsiva. A potência decai rapidamente com o tempo, para uma faixa de corrente superior a 100 mA, demonstrando assim a incapacidade do laser utilizado no experimento em manter a potência em uma faixa ativa de saída.

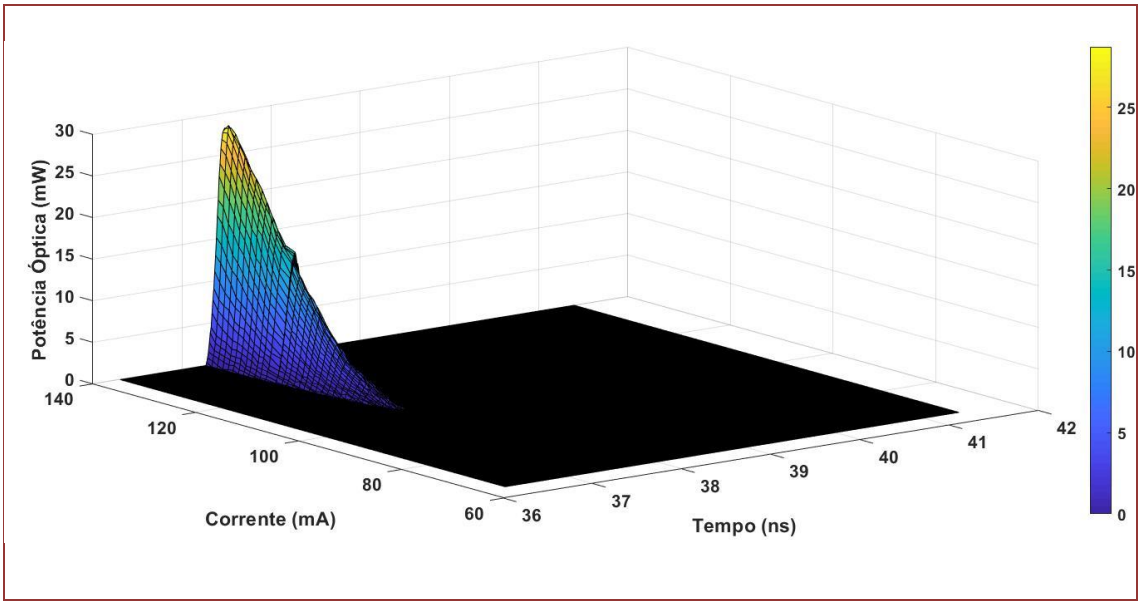


Figura 5.2 - POS (ES à TO = 20ºC). Fonte: Próprio Autor.

Diferentemente da Figura 5.2, na Figura 5.3, apesar de possuir a mesma TO à 20ºC, por ser a saída do sistema para o GS, a potência cresce ao longo do tempo até atingir um valor final na qual estabiliza. Para alguns valores de corrente o comportamento se assemelha ao de um sistema de segunda ordem (LATHI, 2007), melhor ilustrado na Figura 5.4

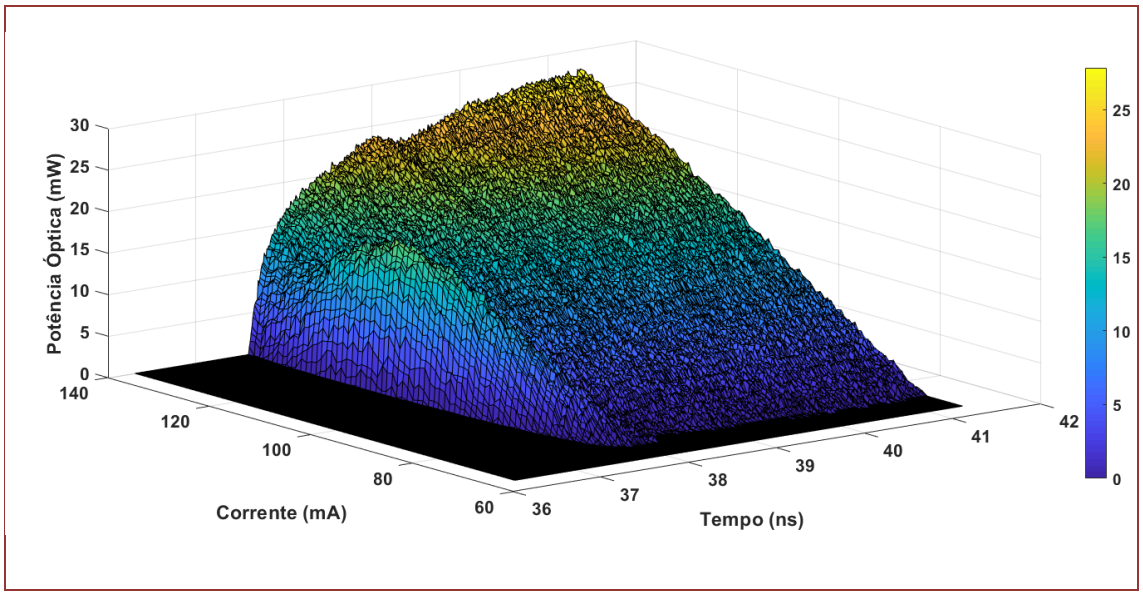


Figura 5.3 - POS (GS à TO = 20ºC). Fonte: Próprio Autor.

A Figura 5.4 ilustra as STs da POS para valores específicos de corrente, tanto para o GS quanto o ES, à TO de 20ºC, pertencentes às amostras apresentadas nas Figuras 5.2 e 5.3, destacando a diferença de comportamento entre os estados para uma mesma corrente de injeção.

Na Figura 5.4 (a) é possível constatar a saída ainda nula do laser com a corrente de injeção de 102 mA no ES, enquanto que no GS a resposta similar de um sistema de segunda ordem super amortecido (OGATA, 2010). Com o aumento da corrente de injeção, é possível visualizar na Figura 5.4 (b) o surgimento do comportamento impulsivo no ES, e a alteração da saída no GS para uma resposta similar a um sistema de segunda ordem sub-amortecido cujo fator de amortecimento próximo à 1 (OGATA, 2010).

Com a alteração na TO, variando de 20ºC para 40ºC, há uma mudança brusca de comportamento, ilustrada na Figura 5.5 para o ES, bem diferente do apresentado na Figura 5.2. Neste caso, já não existe a dificuldade em manter a potência de saída, alterando de uma saída impulsa, para algo similar a uma curva de sistema de segunda ordem, podendo inclusive se observar alguns de seus parâmetros pertinentes (sobre potência, tempo de subida, etc.) nas curvas com correntes mais elevadas.

Já na Figura 5.6, é ilustrado o comportamento do laser no GS a 40ºC, com um comportamento ainda similar a uma curva de um sistema de segunda ordem sub-amortecido, para alguns valores de corrente de injeção, porém com uma suavização que não ocorria a 20ºC.

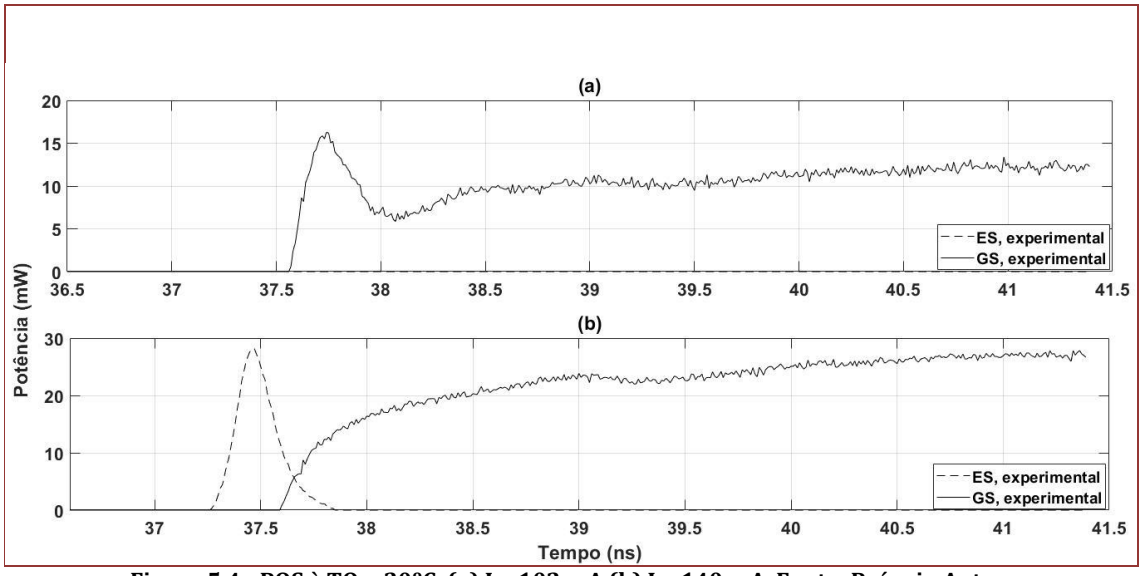


Figura 5.4 - POS à TO = 20ºC. (a) I = 102 mA (b) I = 140 mA. Fonte: Próprio Autor.

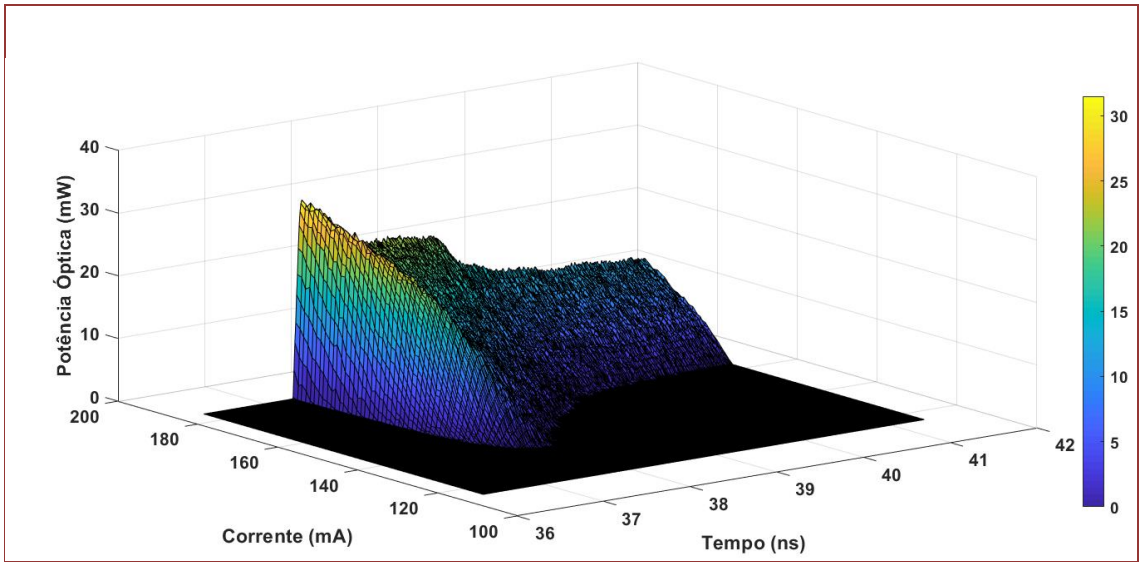


Figura 5.5 - POS (ES à TO = 40ºC). Fonte: Próprio Autor.

Similar à Figura 5.4, a Figura 5.7 ilustra STs e POS para valores específicos de corrente, em ambos estados, à TO de 40ºC, sendo que é possível visualizar o comportamento do laser no ES transitando entre uma resposta impulsiva (Figura 5.7 (a)) para uma resposta similar a um sistema de segunda ordem sub-amortecido

(Figura 5.7 (b)). Simultaneamente, constata-se o comportamento de segunda ordem sub-amortecido do GS, com o aumento do fator de amortecimento com o incremento da corrente de injeção.

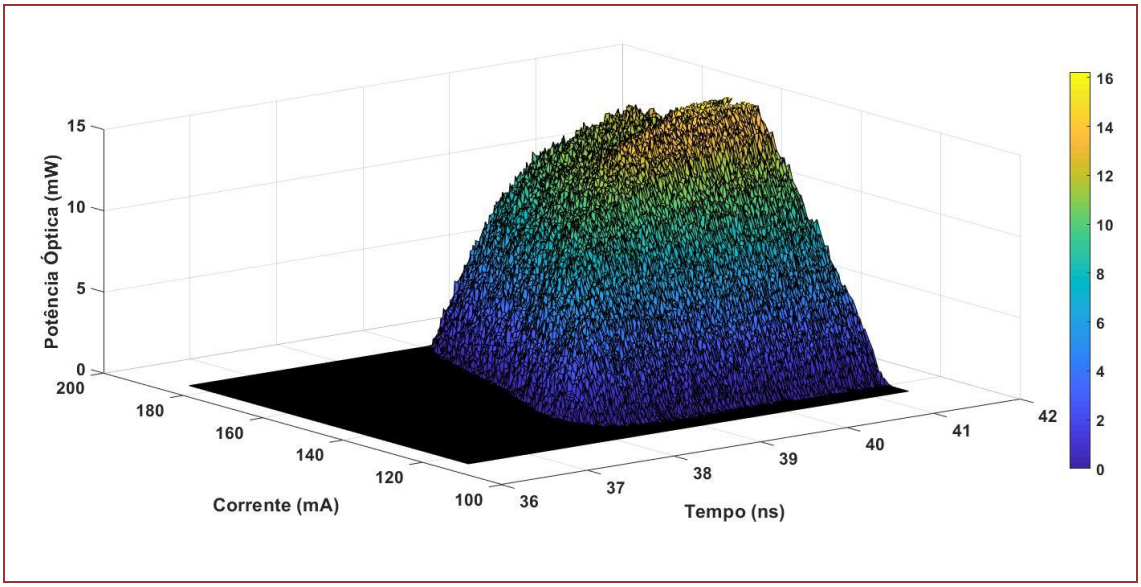


Figura 5.6 - POS (GS à TO = 40ºC). Fonte: Próprio Autor.

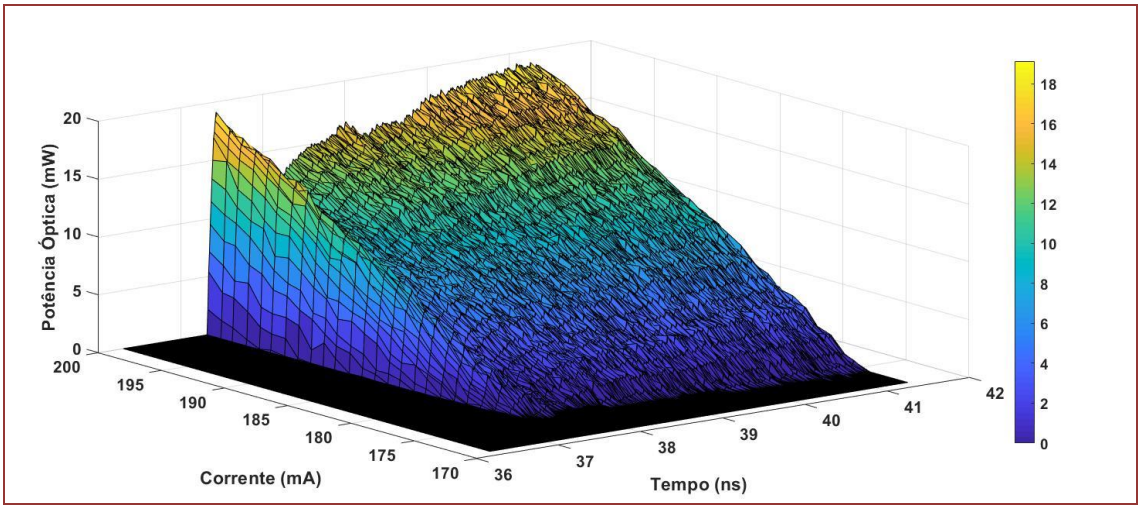


Figura 5.7 - POS à TO = 40ºC. (a) I = 148 mA (b) I = 182 mA. Fonte: Próprio Autor.

No caso específico da TO à 50ºC, o laser apresenta apenas o ES, ilustrado na Figura 5.8, cujo comportamento difere do mesmo estado na temperatura de 40ºC devido a tendência em retornar ao valor de pico após a curva de declínio.

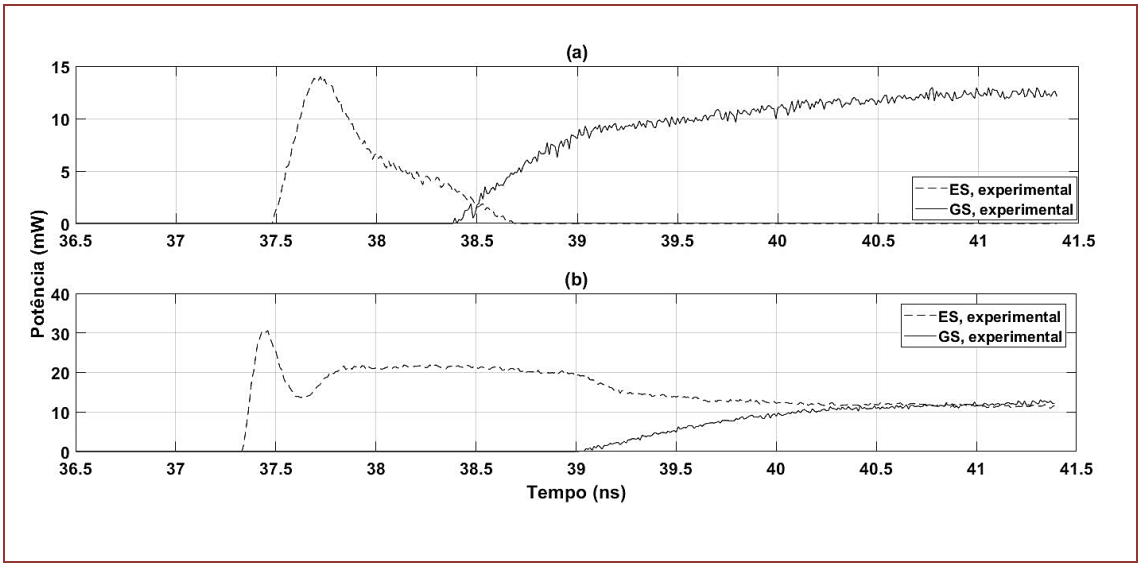


Figura 5.8 - POS (ES à TO = 50ºC). Fonte: Próprio Autor.

A Figura 5.9 destaca duas STs das POS no ES à TO 50ºC, mostrando uma resposta similar a um sistema de segunda ordem sub-amortecido em ambas correntes de injeção (184 e 198 mA), mostrando o aumento dos valores de POS com o aumento da corrente.

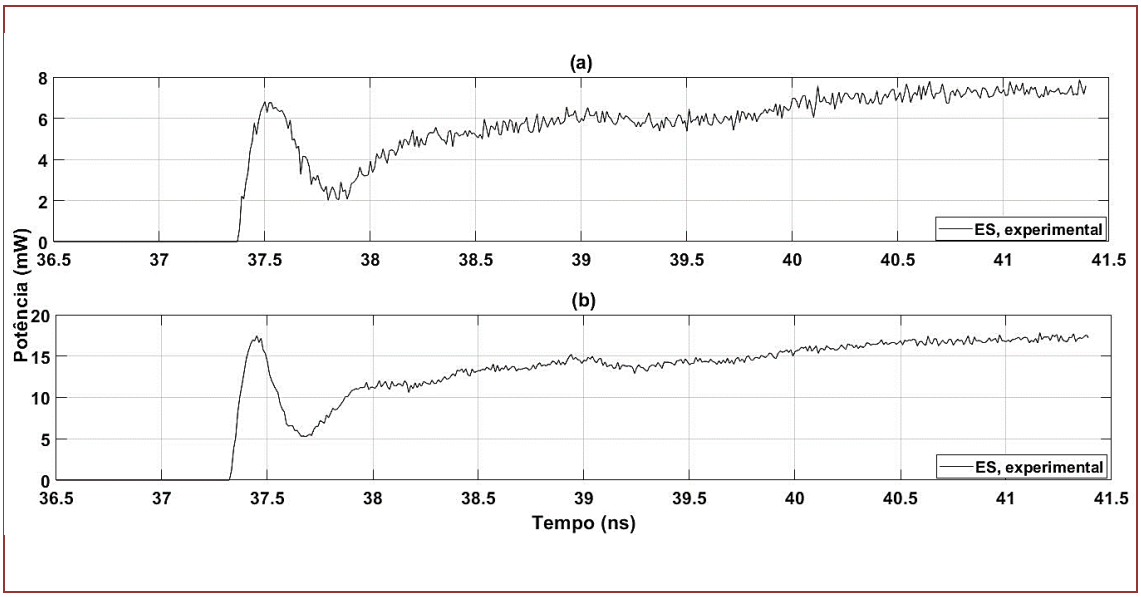


Figura 5.9 - POS à TO = 50ºC. (a) I = 184 mA (b) I = 198 mA. Fonte: Próprio Autor.

5.3.IDENTIFICAÇÃO DO SISTEMA DINÂMICO DO LASER DE PONTOS QUÂNTICOS

5.4.Entradas e Sinal de Excitação

Conforme citado no capítulo 2 deste trabalho, a primeira etapa na identificação do sistema dinâmico, e assim sua modelagem, é a determinação do modelo de entrada, que neste trabalho foi realizado considerando todas as entradas possíveis, que são:

1. Corrente de Excitação (mA), um degrau com corrente entre I_{bottom} e I_{top} detalhados na Tabela 5.1 e ilustradas na Figura 5.10. Devido à natureza das amostras experimentais, a corrente também é o sinal de excitação do sistema;
2. Temperatura de Operação (TO) (ºC), detalhada na Tabela 5.2;
3. Estado de operação do laser, detalhado na Tabela 5.2;
4. Tempo (ns).

Para o melhor desempenho do algoritmo, evitando distorções devido á maior influência no treinamento de determinadas variáveis de entrada com maior faixa dinâmica de valores em relação as demais, todas as amostras de entrada de aprendizagem e teste são normalizadas para o intervalo [0;1] (POS não foi normalizado), usando a seguinte equação:

$$x_{norm}(n) = \frac{x(n)}{x_{máx.}} \tag{5.1}$$

sendo $x_{norm}(n)$ a amostra normalizada, $x(n)$ a amostra original e $x_{máx.}$ o valor máximo do parâmetro.

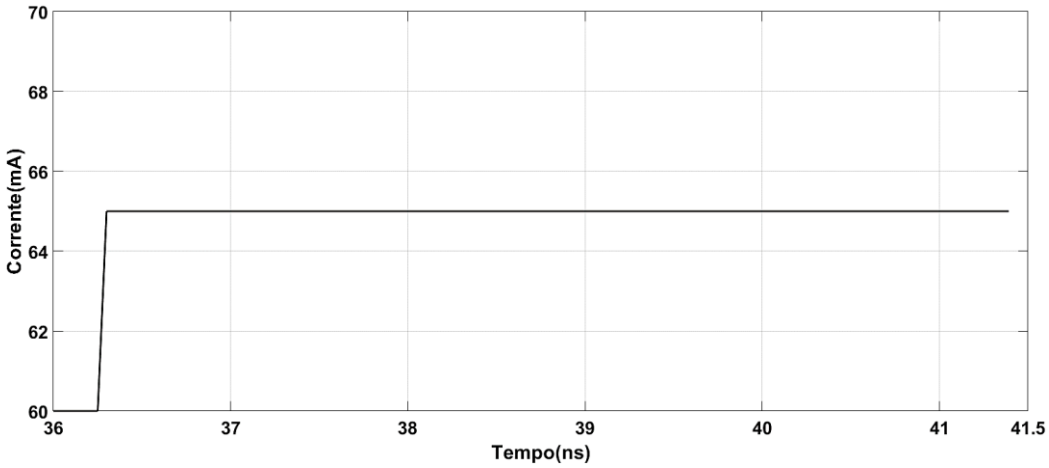


Figura 5.10 - Corrente de Injeção com I_{top} = 65mA. Fonte: Próprio Autor.

De modo similar à normalização das amostras da série temporal, foram padronizados valores no intervalo [0;1] para determinar características intrínsecas das amostras a serem modeladas, e.g., na entrada correspondente ao estado, existem duas opções, 0 para o ES e 1 para o GS, e o mesmo ocorre em relação à temperatura, como detalhado na Tabela 5.2.

Tabela 5.2 – Padronização de Parâmetros de Entrada.

ESTADO	
Descrição	**Valor de Entrada**
ES	0,00
GS	1,00
TEMPERATURA (ºC)	
Descrição	**Valor de Entrada**
20	0,00
40	0,50
50	1,00

5.5.Arquitetura

Na terceira etapa, a escolha da arquitetura do modelo, foi realizada considerando os seguintes critérios:

a) **Tipo de problema**: identificação de sistemas dinâmicos;

b) **Aplicação pretendida**: predição;

c) **Dimensionalidade**: cada modelo proposto possui no mínimo 4 entradas e 1 saída;

d) **Quantidade e qualidade dos dados**: os dados experimentais disponíveis, por serem o resultado de um filtro óptico, com interferências diversas, ocorreram diversos casos com potência ótica com valores negativos, que foram considerados neste trabalho para evitar maiores distorções no modelo.

5.6.Representação Matemática

Na quarta fase, a de representação matemática, foi determinado baseado no conhecimento prévio do experimento e da natureza do laser de pontos quânticos (LPQ), um modelo inicial sem valores autoregressivos, ilustrado na Figura 5.11, e o modelo NARX-RNA, ilustrado na Figura 5.12, cuja função não linear será desempenhada por uma rede neural, ambos com as seguintes características gerais:

a) **Dinâmico:** a POS se altera ao longo do tempo (discreto), assim, utiliza-se equações diferenças na modelagem;

b) **Discretos:** todos os valores de saída foram amostrados;

c) **Multivariáveis:** Foi utilizado o modelo **MISO** em detrimento do modelo MIMO, que possibilitaria prever a saída do laser para ambos os estados simultaneamente, para diminuir a complexidade do modelo a ser proposto;

d) **Paramétricos:** saídas e parâmetros numéricos.

Assim, o modelo apresentado na Figura 5.11 pode ser descrito analiticamente usando:

$$\widehat{POS}(k) = fRNA_{\emptyset}(I, TO, E\ e\ T(k)) \tag{5.2}$$

sendo $\widehat{POS}(k)$ a Potência Óptica de Saída simulada pelo modelo, $fRNA_{\emptyset}(\cdot)$ a função não linear desempenhada pela RNA, e $I, TO, E\ e\ T(k)$ a Corrente de Entrada, a Temperatura de Operação, o Estado e o Tempo no instante k, respectivamente.

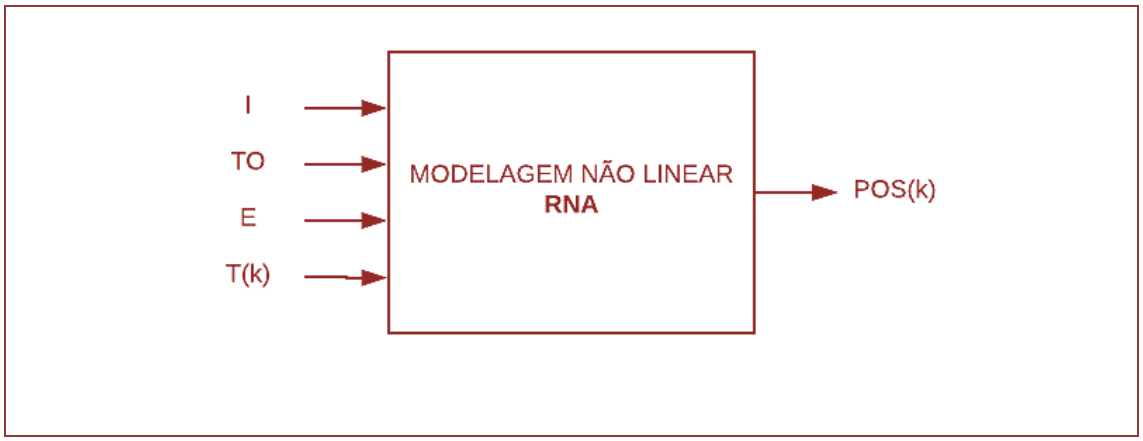


Figura 5.11 - Modelo Não Linear baseado em RNA. Fonte: Próprio Autor.

Já o modelo ilustrado na Figura 5.12 pode ser definido da seguinte maneira:

$$\widehat{POS}(k) = fRNA_{\emptyset}(I, TO, E, T(s)\ e\ POS(p)) \quad (5.3)$$

sendo $s = (k, k-1, k-2, \dots, k-m)$, $p = (k, k-1, k-2, \dots, k-n)$, m o valor de atrasos da entrada, n o valor de atrasos da saída e POS a Potência Óptica de Saída do experimento, já que o tipo de modelagem NARX utilizada neste trabalho foi o série-paralelo (AGUIRRE, 2015).

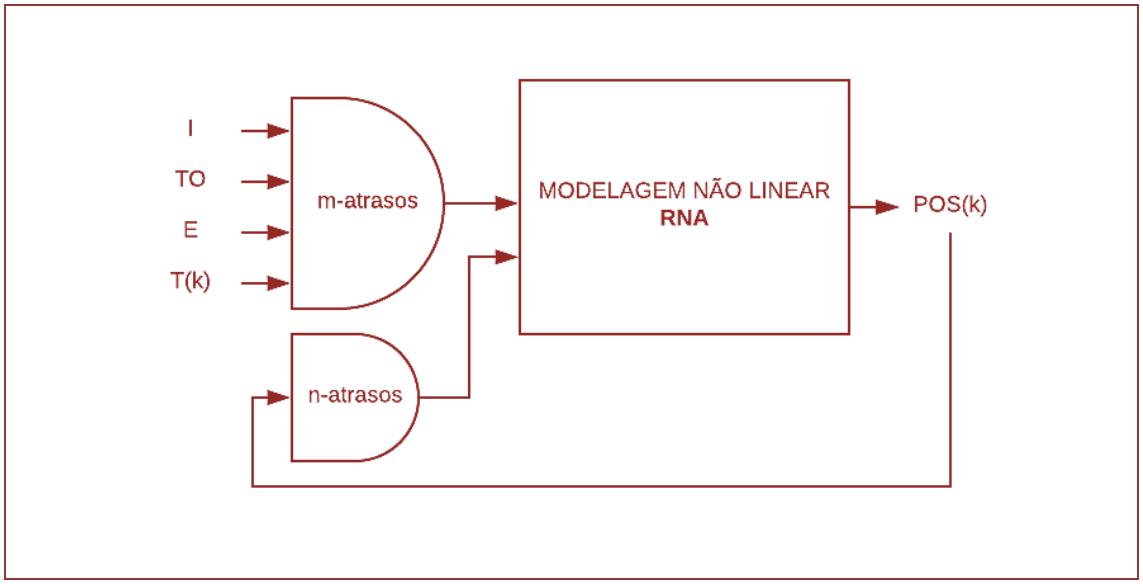


Figura 5.12 – Modelo NARX-RNA. Fonte: Próprio Autor.

Na primeira modelagem, foram utilizadas as seguintes RNAs: ELM, ESN e MLP, além das variantes ABC-ESN e PSO-ESN, na tentativa de otimizar os resultados obtidos para a ESN. Enquanto que a segunda modelagem foi utilizada as mesmas RNAs, e adicionalmente, visando melhorar o desempenho da ESN, foi utilizada a RNA híbrida ESN-ELM.

E para fins de simplificação, a Tabela 5.3 relaciona a nomenclatura adotada para cada modelagem especialista com seu respectivo par estado/temperatura de operação:

Tabela 5.3 - Modelagens Especialistas.

Modelagem Especialista	Estado	Temperatura (ºC)
01	ES	20
02	GS	20
03	ES	40
04	GS	40
05	ES	50

5.7. Ordem, Estrutura e Complexidade do Modelo NARX

Visando uma estrutura mais simples e de menor complexidade, impactando diretamente no tempo de treinamento, limitou-se o aumento da complexidade do modelo NARX até a 5º ordem, priorizando sempre que possível os modelos mais simples.

Para o procedimento de determinação da ordem do NARX, foi utilizado a configuração ótima das RNAs (número de neurônios, função de ativação, etc.) no modelo sem a autoregressão, variando a ordem entre o grau 1 até o grau 5, ou seja, utilizando entre 1 e 5 entradas e saídas defasadas no tempo na entrada da modelagem, sendo cada configuração simulada 100 vezes.

5.8. Parâmetros do Modelo

Similar a música “À procura da batida perfeita” (CORCOS e D2, 2003), onde o cantor procura através dos diferentes estilos musicais seu objeto de perfeição, a determinação de um modelo de predição é uma tarefa fundamental para o sucesso deste trabalho, e perpassa pela implementação de modelos distintos até se alcançar o ideal almejado. Assim, a presente dissertação propõe uma heurística que realiza uma busca por parâmetros de um determinado modelo, originando uma configuração ótima capaz de resolver o problema tema.

Para determinar a configuração ótima deve-se escolher adequadamente os valores das variáveis dos modelos. As variáveis dos modelos propostos são:

- Quantidade de dados experimentais destinados ao treinamento e ao teste;
- Determinação da RNA a ser utilizada;
- Determinação da função de treinamento da RNA;
- Quantidade de neurônios da camada oculta da RNA;
- Determinação da função de ativação da RNA;
- Determinação de algoritmo de otimização;
- Populações a serem utilizadas no algoritmo de otimização.

Apesar da heurística implementar uma busca exaustiva, determinados parâmetros limitadores foram utilizados para determinação de um espaço de soluções plausíveis, diminuindo o custo computacional, que cresce exponencialmente com o aumento de possíveis soluções a serem testadas.

O fluxograma da Figura 5.13 ilustra os passos adotados na heurística deste trabalho, bem como o espaço de soluções possíveis limitados, cujo objetivo é desenvolver um processo que implica em variar alguns parâmetros, enquanto se monitora o comportamento dos índices de desempenho (principalmente o RMSE).

Um dos parâmetros a ser determinado é o algoritmo de treinamento na RNA MLP, possuindo impacto direto no resultado final de sua predição. Assim, para o presente trabalho, foram empregadas as funções apresentadas na Tabela 5.4.

Tabela 5.4 - Funções de Treinamento do MLP.

Algoritmos	
Levenberg-Marquardt (SAPNA, TAMILARASI e KUMA, 2012)	*Fletcher-Powell Conjugate Gradient* (HAGAN, DEMUTH e BEALE, 1996)
Bayesian Regularization (KHAFAF e EL-HAG, 2018)	*Polak-Ribiére Conjugate Gradient* (GHASEMI-FARD, ANSARI-ASL, *et al.*, 2013)
BFGS Quasi-Newton (ZHANG, MU e ZHENG, 2013)	*One Step Secant* (BATTITI, 1992)
Resilient Backpropagation (RIEDMILLER e BRAUN, 1993)	*Gradient descent with momentum and adaptive learning rate backpropagation* (ENDAH, WIDODO, *et al.*, 2017)
Scaled Conjugate Gradient (KHADSE, CHAUDHARI e BORGHATE, 2017)	*Gradient Descent with Momentum* (LENKA e MOHAPATRA, 2015)
Conjugate Gradient with Powell/Beale Restarts (POWELL, 1977)	*Gradient Descent* (HAGAN, DEMUTH e BEALE, 1996)

Em relação as funções de ativação, como detalhado no capítulo anterior, faz-se necessário sua utilização para a implantação do algoritmo da rede neural ELM, influenciando na aleatoriedade inicial dos seus pesos, sendo empregado as seguintes funções de ativação, detalhadas na Tabela 5.5.

Tabela 5.5 - Funções de Ativação.

Função de Ativação	Equação	
Senoidal	$f(x) = sen\ (x)$	(5.4)
Radial Basis	$\emptyset(r) = e^{-(\varepsilon r)^2}$	(5.5)
Triangular Basis	$f(x) = tribas(x) = 1 - \lvert x\rvert, se -1 \leq x \geq 1;$ $senão\ f(x) = 0$	(5.6)
Hard Limit	$f(x) = hardlimit(x) = 1, se\ x \geq 0;$ $senão\ f(x) = -1$	(5.7)
Sigmoide	$f(x) = \frac{1}{1 + e^{-\lambda x}}$	(5.8)

Já na ESN, é necessário parametrizar as variáveis descritas na Tabela 5.6, que variam entre 0 e 1, limitando a conectividade e o raio espectral da rede. Devido à quantidade de variáveis parametrizáveis e faixa de excursão de valores, possibilitando uma elevada gama de configurações, foram utilizados posteriormente os algoritmos de otimização ABC e PSO, visando alcançar a solução ótima da ESN.

Tabela 5.6 - Parâmetros Iniciais da ESN.

Parâmetro	Intervalo	Valores padrão
Probabilidade do peso de entrada	[0;1]	1
Range do vetor de entrada	[0;1]	1
Probabilidade dos pesos recorrentes	[0;1]	0
Range do vetor de pesos recorrentes	[0;1]	0
Probabilidade dos pesos de retorno	[0;1]	0
Range do vetor de pesos retorno	[0;1]	0
Raio espectral	[0;1]	0

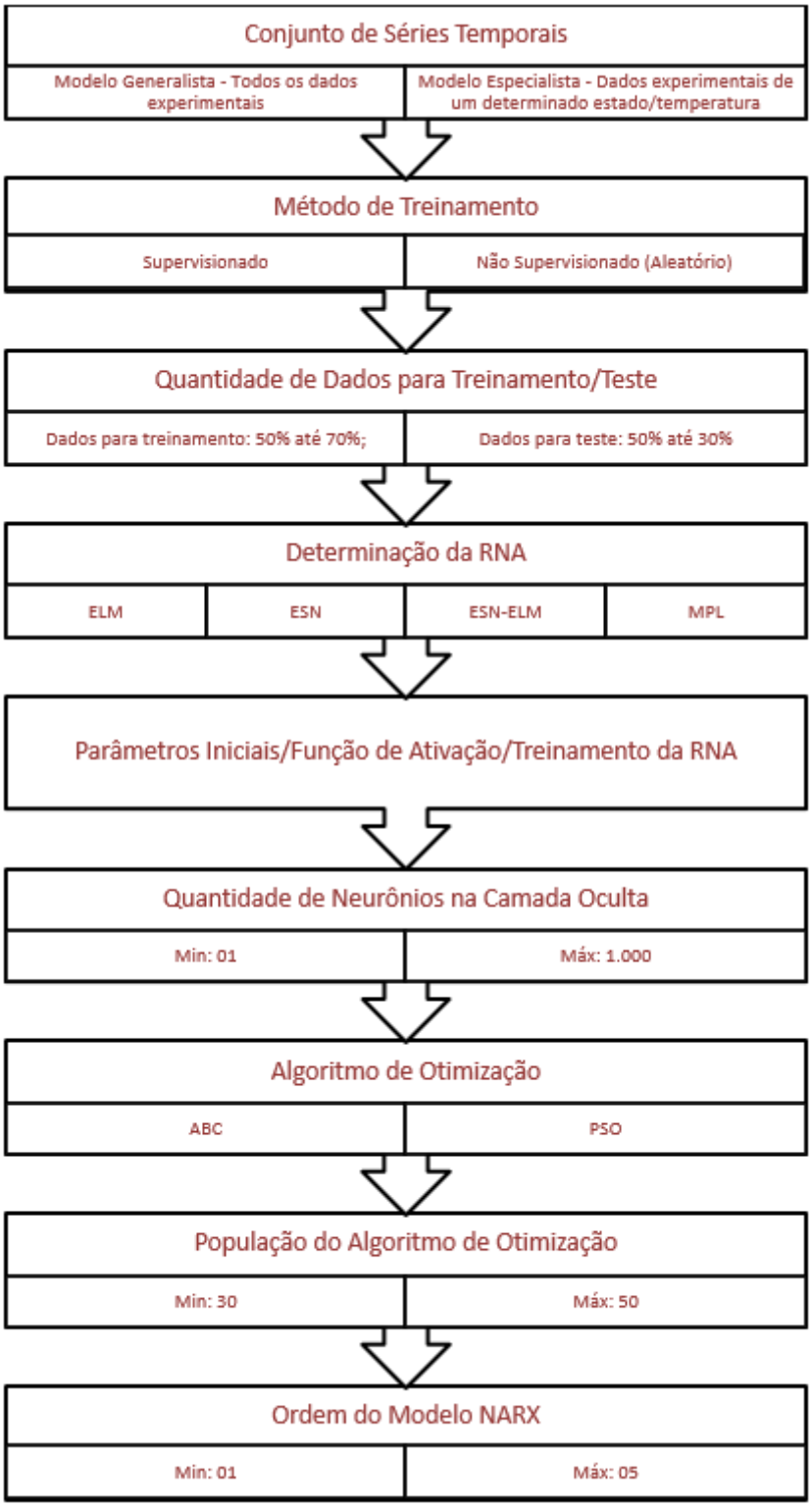


Figura 5.13 - Fluxograma da Heurística. Fonte: Próprio Autor.

5.9.Avaliação de Desempenho

Segundo Menezes Junior (2012), uma importante tarefa em predição de séries temporais é a de quantificar a qualidade dos resultados obtidos, permitindo comparar os diversos algoritmos e estruturas de modelos utilizando índices de desempenho. A seguir são apresentados alguns dos índices de desempenho mais utilizados em avaliação de modelos de regressão, e usados neste trabalho para a validação dos modelos propostos.

A Raiz do Erro Médio Quadrático (*Root Mean Square Error*— RMSE) é definida por:

$$RMSE = \sqrt{\frac{\sum_{t=1}^{n} {e_t}^2}{n}} \tag{5.9}$$

sendo e_t o erro da saída no instante t, que é obtido entre a diferença entre o valor real e o estimado pelo modelo e n é o número de amostras de saída avaliados.

Como consequência, quanto menor o valor RMSE, melhor será o modelo, porém, esta métrica possui a desvantagem de erros menores possuírem menor importância em comparação a erros maiores.

O Erro Percentual Absoluto Médio (*Mean Absolute Percentage Error* — MAPE) é definido por:

$$MAPE = \frac{\sum_{t=1}^{n} |EP_t|}{n} \tag{5.10}$$

sendo EP_t o erro percentual da saída no instante t e n o número de amostras de saída avaliados.

Similar ao RMSE, quanto menor o valor de MAPE, melhor o desempenho do modelo em análise, mas possui uma penalidade maior para as previsões que superestimam os valores reais do que em relação aquelas que os subestimam, criando assim distorções.

Devido as desvantagens citadas acima do RMSE e do MAPE, foi considerado outro importante parâmetro utilizado na avaliação de modelos de regressão, o índice proposto em (THEIL, 1966).

A Estatística U de Theil foi desenvolvida visando avaliar o modelo de regressão comparando-o com o erro da previsão ingênua, que indica que o valor no tempo t será o valor da última observação, ou seja, igual ao valor obtido em $t - 1$.

Qualquer modelo de previsão para séries temporais, linear ou não-linear, deve apresentar como requisito mínimo um desempenho melhor que o da previsão ingênua, ou seja, apresentar U<1, e quanto mais próximo de 0, melhor o modelo de predição.

$$U(Theil) = \frac{\sqrt{\frac{\sum_{t=1}^{n}(Y_t - YP_t)^2}{n}}}{\sqrt{\frac{\sum_{t=1}^{n}(Y_t)^2}{n}} + \sqrt{\frac{\sum_{t=1}^{n}(YP_t)^2}{n}}} \tag{5.11}$$

sendo $U(Theil)$ o valor de Estatística de U de Theil, n a quantidade de amostras, Y_t o valor de saída obtido pelo algoritmo e YP_t a saída da previsão ingênua.

Em paralelo aos índices de desempenho estatísticos, foi-se avaliado o tempo total de processamento do modelo utilizado, definido por:

$$T.T.P. = t_{tr} + t_{ts} \tag{5.12}$$

sendo $T.T.P.$ o tempo total de processamento, t_{tr} o tempo de treinamento e t_{ts} o tempo de teste da modelagem avaliada.

Tal avaliação é fundamental para futuras comparações com modelos tradicionais de equações de estados e como comparativo entre modelos testados, já que se pode escolher um modelo com menor tempo total de processamento em detrimento do seu RMSE superior, e.g., a depender da aplicação e nível de confiabilidade exigido.

5.10. CONSIDERAÇÕES

Este capítulo apresentou as amostras experimentais adquiridas previamente e usados neste trabalho, juntamente com a metodologia utilizada. Inicialmente apresentou-se a fonte e características dos dados e posteriormente todo o procedimento realizado para a identificação do sistema dinâmico, das padronizações a determinação dos parâmetros básicos, a heurística para a determinação do melhor modelo de predição dentre os implantados e os índices de desempenho.

6 Resultados

6.1.INTRODUÇÃO

No atual capítulo são discutidos os resultados obtidos pelos diversos modelos propostos, iniciando pela determinação dos parâmetros de cada técnica utilizada, em seguida a análise dos resultados globais, e finalizando uma análise mais detalhada do melhor modelo obtido na predição do sistema dinâmico de lasers de pontos quânticos.

6.2.RESULTADOS PARA DETERMINAÇÃO DE PARÂMETROS

Seguindo o fluxograma ilustrado Figura 5.13, foram feitas inúmeras simulações, gerando resultados intermediários, até a determinação dos parâmetros responsáveis pelos resultados ótimos de cada modelo. Foi descrito a seguir o procedimento adotado para a determinação dos parâmetros para o modelo generalista ELM. Uma metodologia semelhante foi utilizada para os demais casos.

As Figuras 6.1 e 6.2 ilustram a determinação da função de ativação ideal para um modelo ELM Generalista utilizando-se 70% das amostras para o treinamento e 30% para o teste, variando paralelamente tanto a função de ativação quanto a quantidade de neurônios.

A Figura 6.1 ilustra a variação do RMSE por função de ativação para a ELM Generalista, caindo exponencialmente com o aumento da quantidade de neurônios até 200. Assim, é possível visualizar que o RMSE da função *Hard Limit* é significativamente superior as demais funções, sendo prontamente descartado para futuros testes de modelagens neste trabalho.

Já na Figura 6.2 visualiza-se a variação linear do tempo total de treinamento (tempo em segundos) de cada função de ativação com o aumento do número de neurônios até 200. Como a variação entre os tempos de processamento são

similares, não havendo uma grande discrepância entre os tempos, não foi descartado nenhuma função de ativação por esta avaliação.

Já no procedimento da determinação do número de neurônios da camada oculta do modelo ELM Generalista com função de ativação Senoidal (escolhida dentre as testadas anteriormente), é verificado o menor RMSE de cada quantidade determinada, variando a quantidade de neurônios de 01 até 1.000, como ilustrado na Figura 6.3.

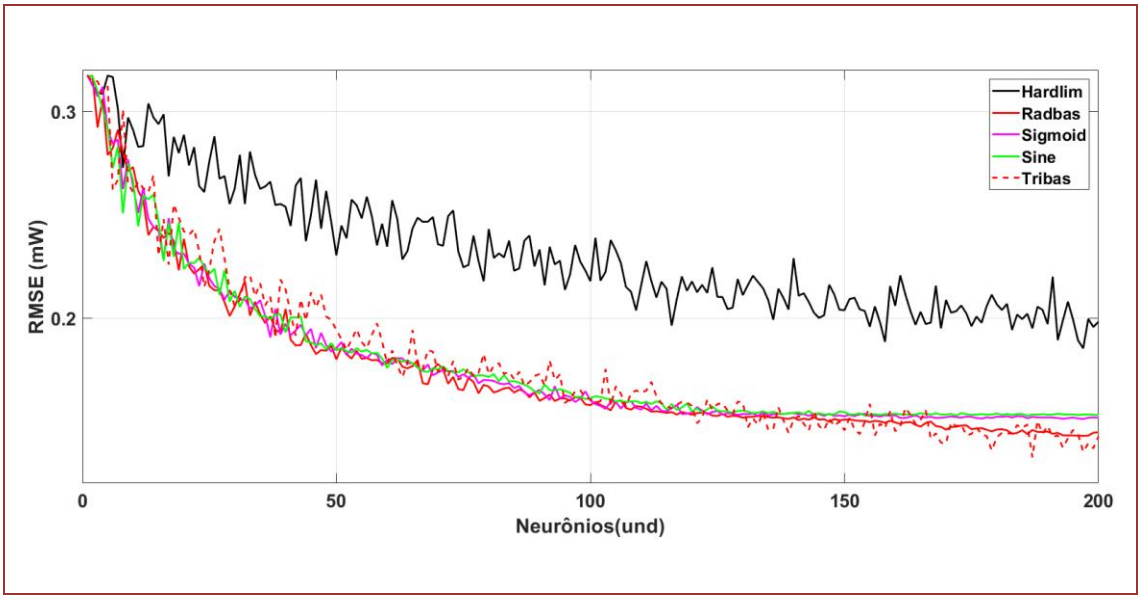


Figura 6.1 - Modelo ELM Generalista - Função de Ativação x Neurônios x RMSE. Fonte: Próprio Autor.

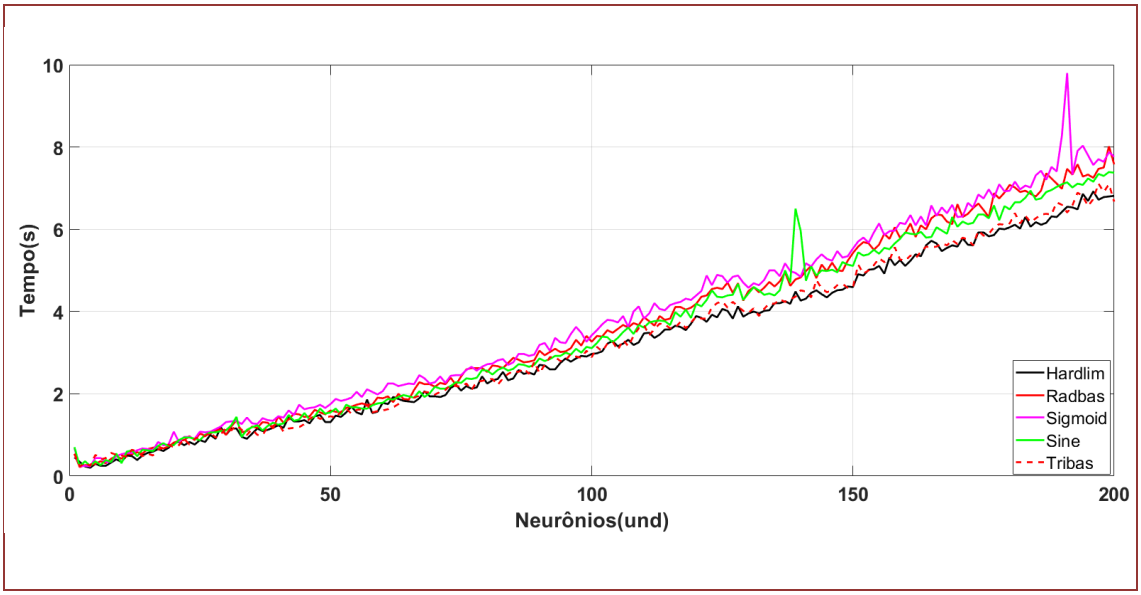


Figura 6.2 - Modelo ELM Generalista - Função de Ativação x Neurônios x Tempo. Fonte: Próprio Autor.

O melhor resultado obtido foi com 552 neurônios na camada oculta, acima desta quantidade verificou-se o aumento do RMSE em detrimento da performance devido a quantidade superior de neurônios. Não é possível visualizar graficamente a quantidade de neurônio com menor RMSE, devido a certa estabilidade gerada e pequena variação entre os resultados após 200 neurônios.

Assim, procedimentos similares foram realizados na determinação dos parâmetros dos demais modelos, com seus parâmetros finais definidos na Tabela 6.5 e resultados detalhados na Tabela 6.4, discutidos ao longo deste capítulo.

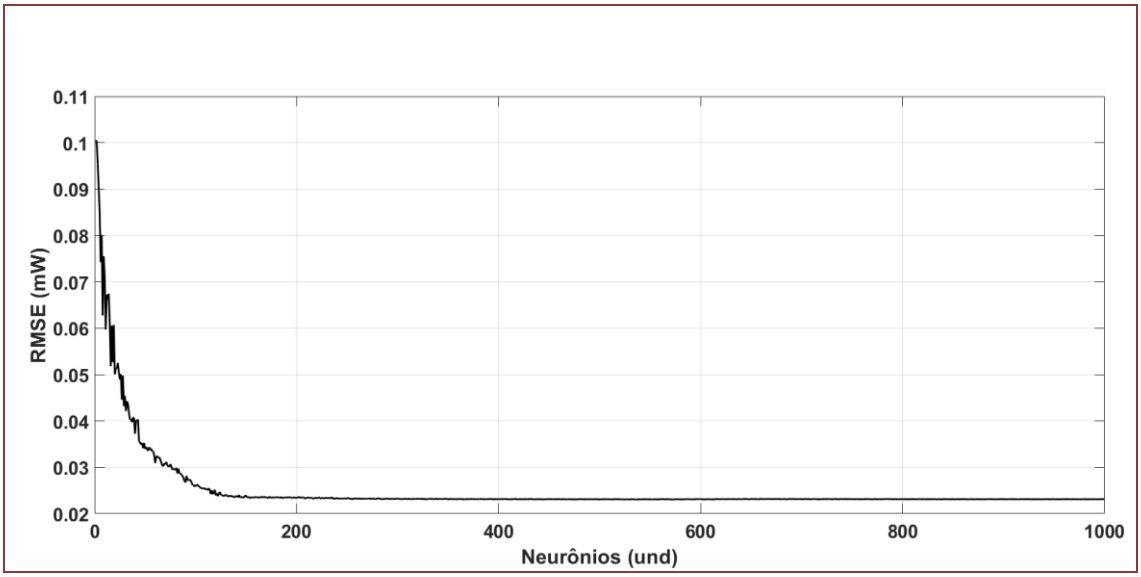


Figura 6.3 - ELM Generalista - RMSE x Neurônios. Fonte: Próprio Autor.

6.3.ANÁLISE DOS RESULTADOS GLOBAIS

Esta seção tem como objetivo avaliar os resultados globais obtidos por cada modelo proposto neste trabalho, assim como os parâmetros necessários para alcançar seus melhores desempenhos.

6.4.Modelos Generalistas

Como ilustrado no Fluxograma da Heurística (Figura 3.2), inicialmente foi avaliado o desempenho das RNAs no modelo generalista, modelo este que utiliza todas as amostras, sem distinção com relação ao par estado/temperatura. Foi verificado a vantagem da MLP em relação ao desempenho, que mesmo nesta abordagem já apresentou resultados satisfatórios (RMSE = 0,33 mW, MAPE = 2,15%

e U = 0,0161), porém em detrimento do tempo total de processamento, superior aos demais, de 21,60 minutos, enquanto a ELM foi de apenas 0,97 minutos e a ESN de 2,81 minutos.

Apesar do seu tempo total de processamento ser de apenas 4,5% do MLP, a ELM apresentou um péssimo resultado em relação as demais (RMSE = 2,30 mW, MAPE = 12,92% e U = 0,1156), demonstrando-se ineficaz na modelagem generalista. De forma similar, a RNA ESN também não obteve sucesso (RMSE = 1,70 mW, MAPE = 9,23% e U = 0,0852), porém com um tempo total de processamento de apenas 2,81 minutos.

Em relação aos parâmetros das RNAs na modelagem generalista, o MLP obteve seu melhor resultado (tanto em relação ao RMSE, quanto o MAPE e o U) com 91 neurônios na camada oculta, com o algoritmo de treinamento de *Levenberg-Marquardt* (SAPNA, TAMILARASI e KUMA, 2012), ocorrendo aumento do RMSE com o incremento do número de neurônios além deste valor, ou seja, um aumento na complexidade da modelagem aquém do necessário, piorando seu desempenho.

No ELM, observou-se uma certa estagnação no desempenho, com uma leve piora com a utilização superior a 552 neurônios na camada oculta, e a ESN ocorreu um crescimento do desempenho com um aumento dos neurônios, o que resultou na utilização da quantidade máxima testada de 1.000 neurônios.

Ambos obtiveram resultados superiores com a função de inicialização senoidal, e as três modelagens tiveram seu melhor desempenho com o treinamento supervisionado na relação 70% treinamento e 30% teste, considerando amostras espaçadas de forma equidistantes nas séries temporais.

Assim, para todos os demais testes de modelos propostos posteriormente, foi padronizado a utilização dos seguintes parâmetros básicos:

- Função de Inicialização do ELM e ESN: senoidal;
- Algoritmo de Treinamento do MLP: *Levenberg-Marquardt;*
- Treinamento supervisionado na relação 70%/30% de amostras para treino/teste.

Na Figura 6.4 é possível visualizar os resultados obtidos pelos modelos generalistas na predição da dinâmica do LPQ no GS à TO de 20ºC para corrente

iguais a 102 e 140 mA, confirmando o desempenho insatisfatório do ELM e da ESN na predição neste modelo, principalmente no regime transitório, enquanto o MLP consegue predizer toda a dinâmica sem apresentar um erro elevado.

Já a Tabela 6.1 apresenta resumidamente os valores alcançados para cada técnica utilizada na modelagem generalista, transparecendo a grande diferença de desempenho (RMSE, MAPE e U) entre a MLP, o ELM e a ESN, apesar de ambos terem o tempo total de processamento muito inferior ao MLP.

Adicionalmente, a Tabela 6.2 apresenta as configurações ótimas de cada técnica utilizada na modelagem generalista.

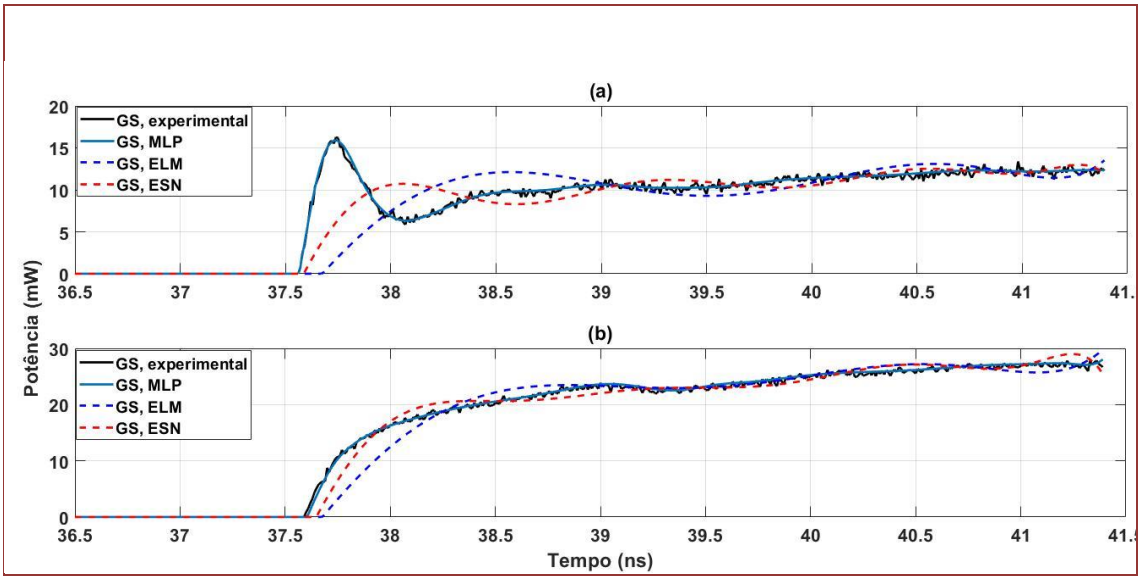


Figura 6.4 - POS Modelo Generalista à TO = 20ºC. (a) I = 102 mA (b) I = 140 mA. Fonte: Próprio Autor.

Tabela 6.1 - Resultados Globais do Modelo Generalista.

Modelo	RMSE (mW)	MAPE (%)	U de Theil (adimensional)	Tempo Total de Processamento (min)
ELM GENERALISTA	2,30	12,92	0,1156	0,97
ESN GENERALISTA	1,70	9,23	0,0852	2,81
MLP GENERALISTA	0,33	2,15	0,0161	21,60

Tabela 6.2 - Configurações dos Parâmetros do Modelo Generalista.

Modelo	Número de Neurônios	Função de inicialização/treinamento
ELM GENERALISTA	552	Senoidal
ESN GENERALISTA	1.000	Senoidal
MLP GENERALISTA	91	Levenberg-Marquardt

6.5.Modelos Especialistas

Devido ao fracasso da modelagem generalista utilizando as RNAs ELM e ESN, retornou-se ao início do Fluxograma da Heurística (Figura 5.13), para a verificação do desempenho das RNAs no modelo especialista, separando os dados de treinamento e testes por subgrupos determinados pelo par estado/temperatura de operação do resultado experimental, ou seja, no modelo especialista existem 5 redes neurais para cada diferente modelagem com suas respectivas características.

Similar ao modelo generalista, a modelagem com MLP obteve o melhor desempenho (RMSE = 0,28 mW, MAPE = 1,51% e U = 0,0139), porém com um tempo total de processamento ainda mais elevado, de 76,30 minutos. Comparando ambas modelagens com MLP, apesar da melhora do desempenho no modelo especialista, este não foi significativo, principalmente levando-se em conta o aumento no tempo total de processamento em 353%, justificada pela quantidade de RNAs MLP no modelo ser superior ao generalista.

No caso da modelagem ELM especialista, houve uma melhora no desempenho (RMSE = 2,22 mW, MAPE = 10,98% e U =0,1115), apesar de inferior as demais RNAs, e uma diminuição do tempo total de processamento em cerca de 55%, devido a simplificação da modelagem com a divisão dos dados em subgrupos, diminuindo assim a quantidade de neurônios necessárias para seu melhor desempenho.

Na modelagem com a ESN, houve uma melhora do desempenho, não significativa, quando se visualiza a métrica MAPE (de 9,23% para 8,83%), porém uma piora em relação ao RMSE (de 1,70 mW para 15,72 mW), além do aumento do tempo total de processamento para 3,10 minutos.

A Figura 6.5 ilustra os resultados obtidos pelos modelos especialistas, retratando o insatisfatório desempenho do ELM e da ESN que está condicionado à uma escolha ótima dos parâmetros, assim, no intuito de melhorar o desempenho da ESN, foram utilizados algoritmos de otimização, o ABC e o PSO, para maximizar seu desempenho alterando sua parametrização inicial.

Entretanto, apesar do excessivo tempo total de processamento gerado pelos algoritmos de otimização (ABC com 21.887 minutos e o PSO com 9.520 minutos), os resultados ficaram muito aquém do esperado, demonstrando inclusive pioras em relação a modelagem ESN padrão, como pode ser visualizado na Tabela 6.4. Um dos motivos da piora destes resultados em detrimento da sua otimização é o excessivo número de parâmetros escolhidos para otimização pelos algoritmos, a grande faixa dinâmica de possíveis soluções, além da dificuldade dos algoritmos em relação a mínimos locais, sendo assim, descartados das modelagens posteriores.

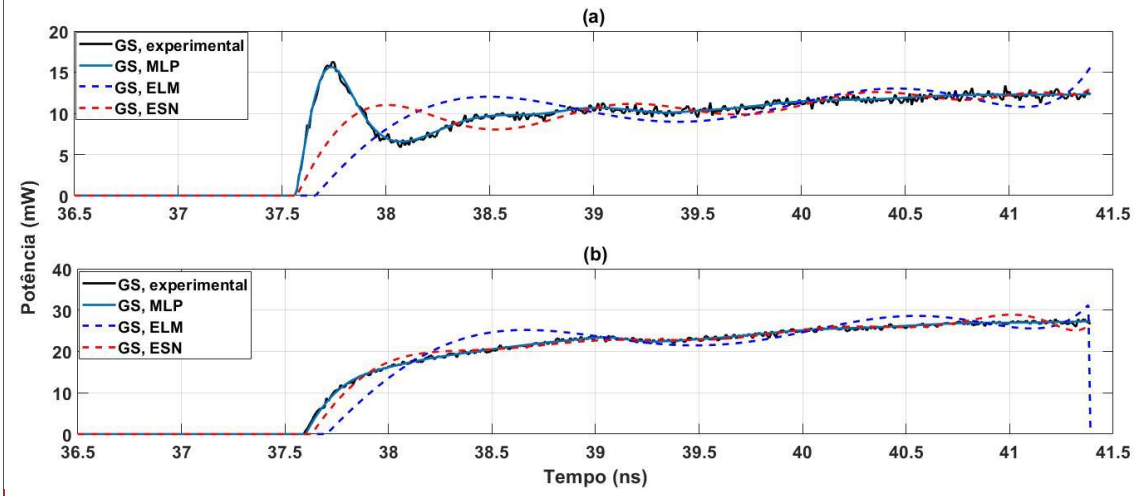


Figura 6.5 - POS Modelo Especialista à TO = 20ºC. (a) I = 102 mA (b) I = 140 mA. Fonte: Próprio Autor.

As Figuras 6.6 e 6.7 ilustram os resultados dos modelos especialistas ESN com os algoritmos de otimização, apresentando graficamente o péssimo resultado obtido com o PSO-ESN. Entretanto, o ABC-ESN obteve sucesso na modelagem do LPQ no GS à TO de 20ºC, superando o resultado do modelo ESN padrão, porém apresentou um resultado inferior ao se comparar ao mesmo no ES à TO de 40ºC, como ilustrado na Figura 6.7.

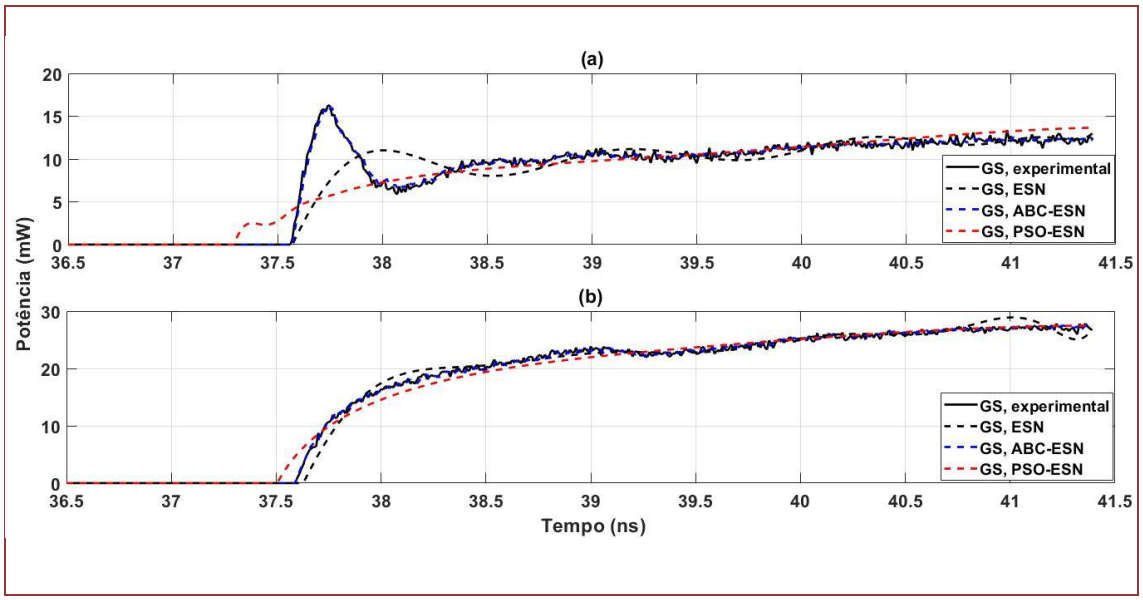


Figura 6.6 - POS Modelo Especialista SOA à TO = 20ºC. (a) I = 102 mA (b) I = 140 mA. Fonte: Próprio Autor.

A Tabela 6.3 apresenta resumidamente os valores alcançados para cada técnica utilizada na modelagem especialista, onde visualiza-se o excelente desempenho (RMSE, MAPE e U) da MLP, e os resultados insatisfatórios do ELM e a ESN, mesmo após a separação das amostras por par estado/temperatura, e o fracasso na melhora do desempenho da ESN com a utilização dos algoritmos ABC e PSO.

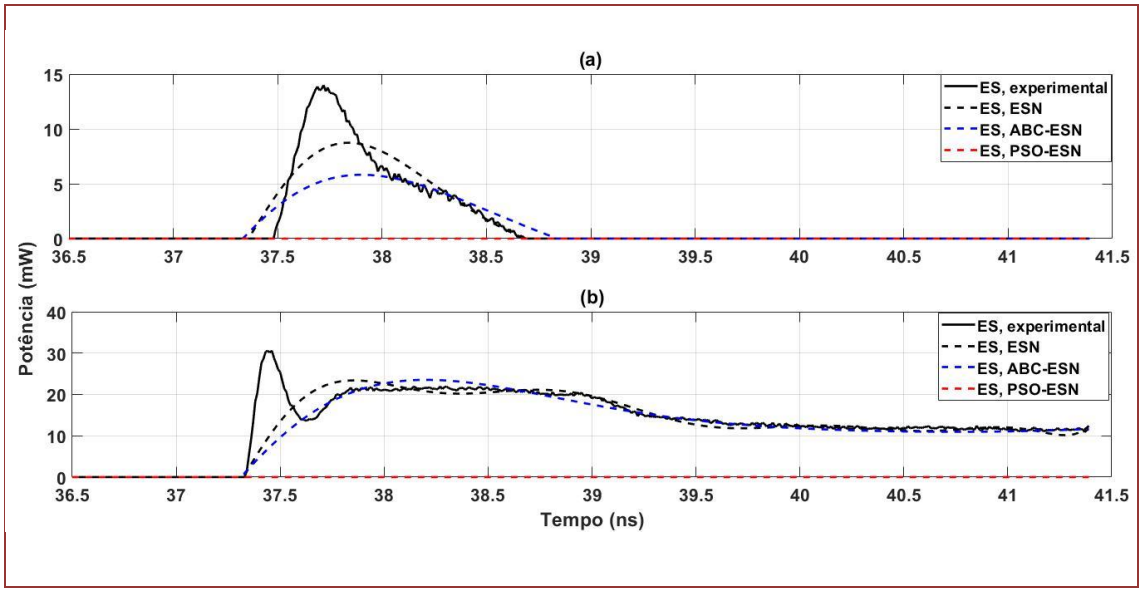


Figura 6.7 - POS Modelo Especialista SOA à TO = 40ºC. (a) I = 148 mA (b) I = 182 mA. Fonte: Próprio Autor.

Tabela 6.3 - Resultados Globais do Modelo Especialista.

Modelo	RMSE (mW)	MAPE (%)	U de Theil (adimensional)	Tempo Total de Processamento (min)
ELM ESPECIALISTA	2,22	10,98	0,1115	0,43
ESN ESPECIALISTA	15,72	8,83	0,0784	3,10
ABC+ESN ESPECIALISTA	1,97	9,11	0,0982	21.887,99
PSO+ESN ESPECIALISTA	50,37	186,49	0,8413	9.520,40
MLP ESPECIALISTA	0,28	1,51	0,0139	76,30

6.6.NARX-Especialistas

Continuando o Fluxograma da Heurística (Figura 5.13), foi utilizado a modelagem especialista NARX + RNAs, como último passo, onde novamente o melhor desempenho alcançado foi com o modelo NARX-MLP especialista de 4º ordem (RMSE = 0,29 mW, MAPE = 2,09% e U = 0,0141), porém inferior ao modelo especialista proposto anteriormente, além de um tempo total de processamento maior de 161,25, ou seja, dos modelos propostos com a RNA MLP, o melhor resultado obtido foi com o Modelo Especialista.

O NARX-ELM especialista de 1º ordem obteve o pior resultado dentre todos os modelos proposto com a RNA ELM (RMSE = 2,70 mW, MAPE = 15,09% e U = 0,1365), além do aumento no tempo total de processamento, para 1,55 minutos, consequência da necessidade de um número maior de neurônios, como é possível visualizar na Tabela 6.5, sendo assim, atestado a inviabilidade da utilização da RNA ELM para a predição da dinâmica de LPQs.

Analogamente, o NARX-ESN especialista de 1º ordem teve um péssimo resultado (RMSE = 1,19 mW, MAPE = 15,29% e U = 0,0593), entretanto com um tempo total de processamento inferior a todos os demais modelos propostos, de apenas 0,04 minutos, consequência da quantidade extremamente inferior de neurônios utilizados em comparação aos demais.

Assim, como uma tentativa de alcançar um melhor resultado com a ESN e o ELM, principalmente devido ao tempo total de processamento muito inferior se comparado aos modelos com MLP, foi proposto um NARX-ESN-ELM especialista de 1º ordem, utilizando a RNA híbrida apresentada anteriormente, o que resultou em um excelente desempenho (RMSE = 0,37 mW, MAPE = 3,79% e U = 0,0181), similar aos resultados obtidos com modelos MLP, com a vantagem de um tempo total de processamento inferior, de apenas 0,09 minuto, assim, sendo escolhido como o melhor modelos entre os propostos para a predição do LPQ, e sendo avaliado mais detalhadamente na seção a seguir.

A Figura 6.8 ilustra os resultados com os modelos NARX descritos anteriormente, para duas determinadas STs, onde se observa a piora do resultado do MLP frente às simulações anteriores com o mesmo e o comportamento insatisfatório do ELM e da ESN. Porém, é visível que a utilização da rede híbrida ESN-ELM apresentou um ótimo resultado em ambas series temporais ilustradas, sem nenhum erro de predição substancial ao longo de toda a predição.

Já as Figuras 6.9 até 6.18 ilustram a POS das amostras experimentas e todos os resultados obtidos com o modelo NARX-ESN-ELM Especialista, sendo que cada figura ilustrando os dados para um determinado par estado/temperatura. Comparando-se o resultado e as amostras, é visível o sucesso da predição com este modelo para todas as condições determinadas no presente trabalho.

Adicionalmente, a Tabela 6.4 apresenta todos os resultados globais dos modelos analisado, onde cada linha descreve os resultados obtidos para a modelagem para um par temperatura/estado, inclusive os resultados apresentados anteriormente, objetivando em uma maior facilidade de comparação de desempenho, com seus parâmetros ótimo considerados apresentados na Tabela 6.5.

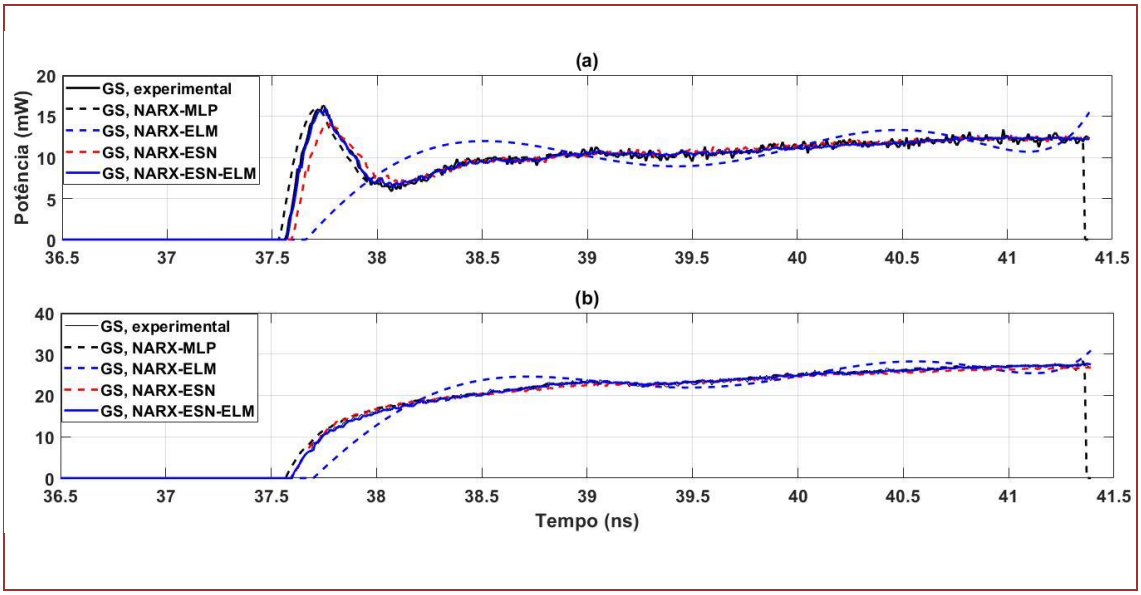


Figura 6.8 - POS Modelo NARX Especialista à TO = 20ºC. (a) I = 102 mA (b) I = 140 mA. Fonte: Próprio Autor.

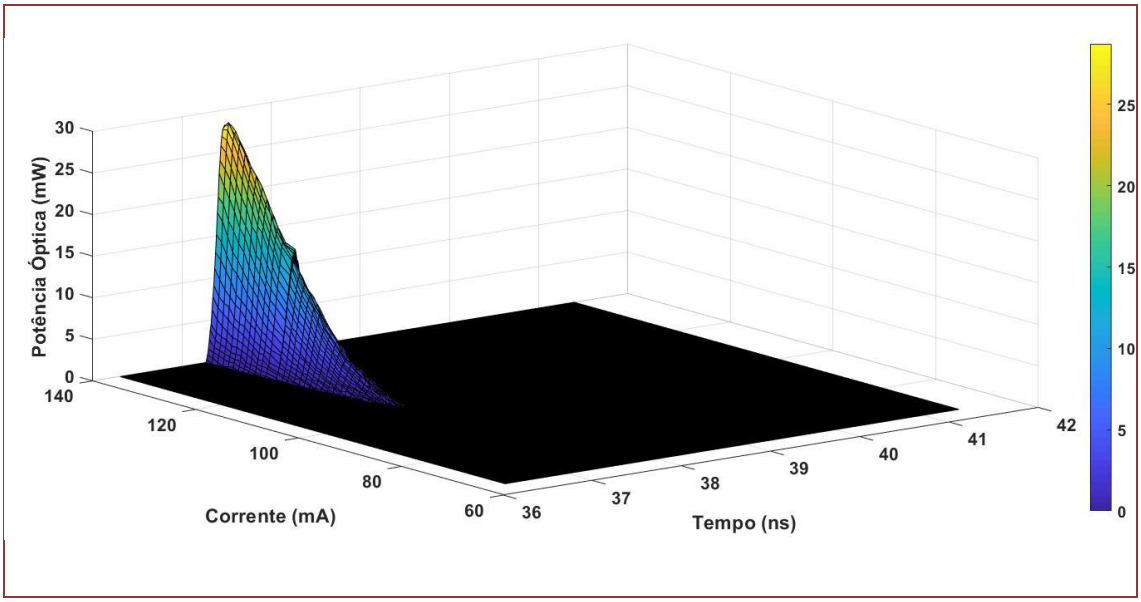


Figura 6.9 – POS (ES à TO = 20ºC). Fonte: Próprio Autor.

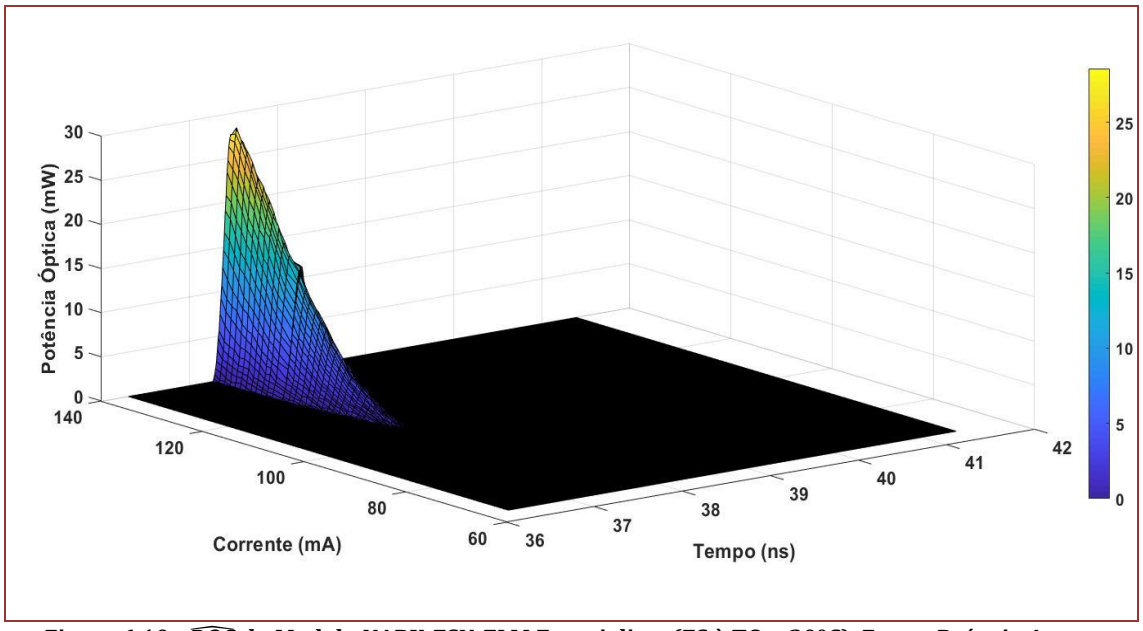


Figura 6.10 - $\widehat{POS}$ do Modelo NARX-ESN-ELM Especialista (ES à TO = 20ºC). Fonte: Próprio Autor.

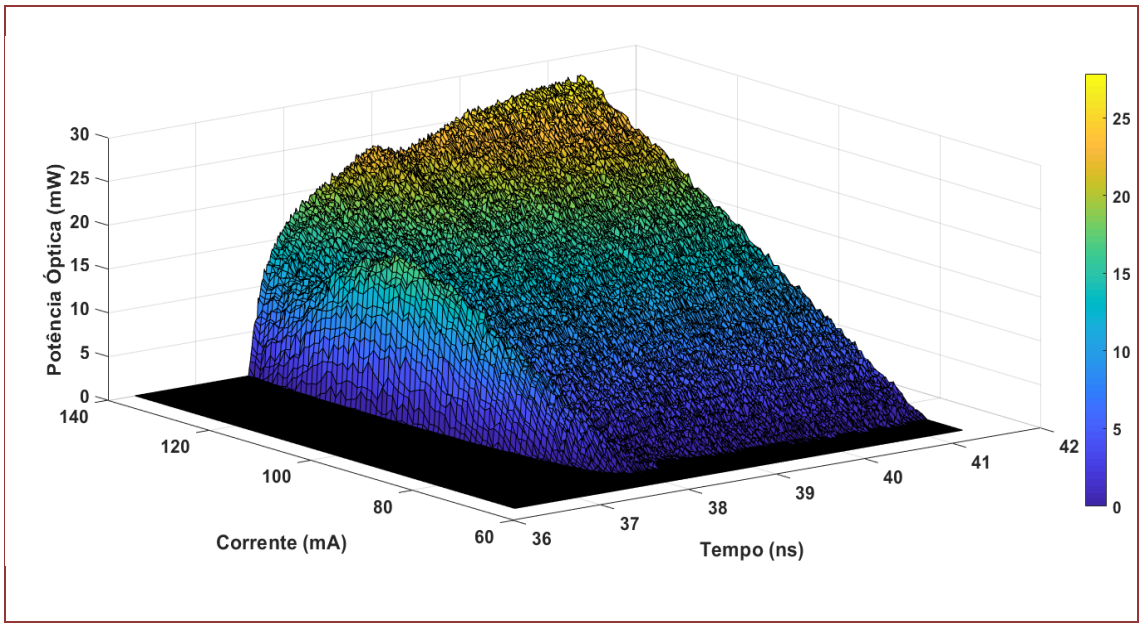


Figura 6.11 – POS (GS à TO = 20ºC). Fonte: Próprio Autor.

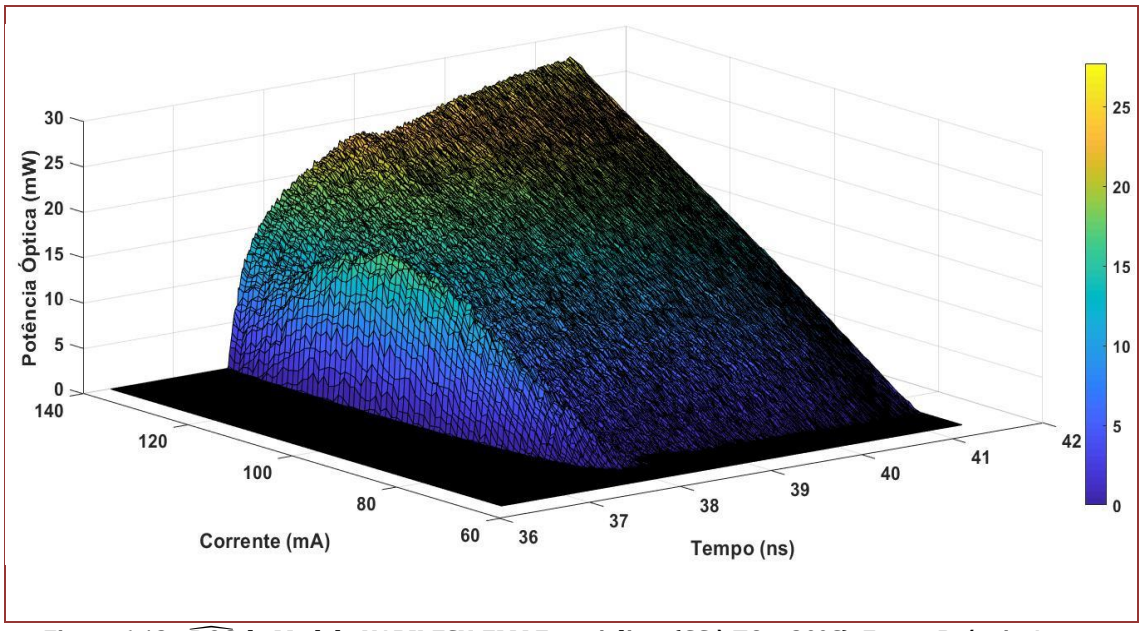


Figura 6.12 - $\widehat{POS}$ do Modelo NARX-ESN-ELM Especialista (GS à TO = 20ºC). Fonte: Próprio Autor.

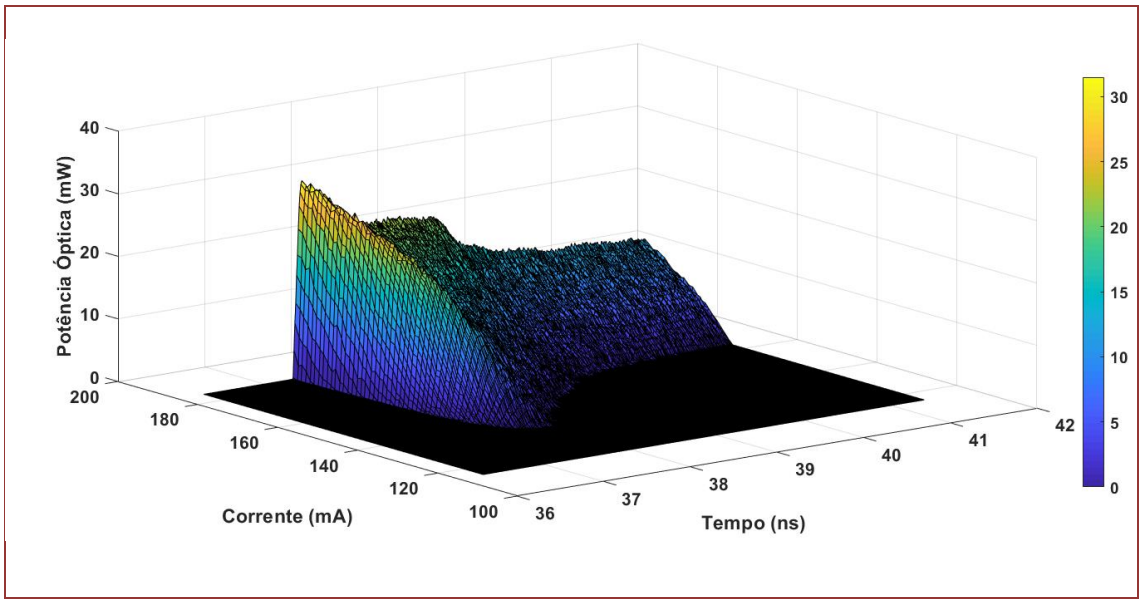


Figura 6.13 - POS (ES à TO = 40ºC). Fonte: Próprio Autor.

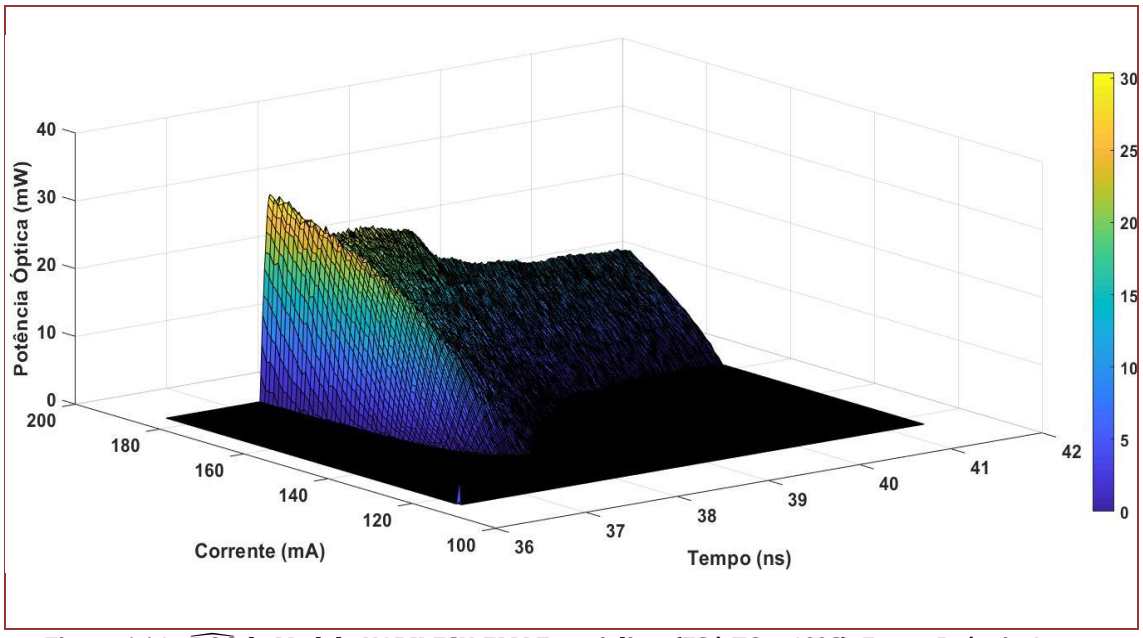


Figura 6.14 - $\widehat{POS}$ do Modelo NARX-ESN-ELM Especialista (ES à TO = 40ºC). Fonte: Próprio Autor.

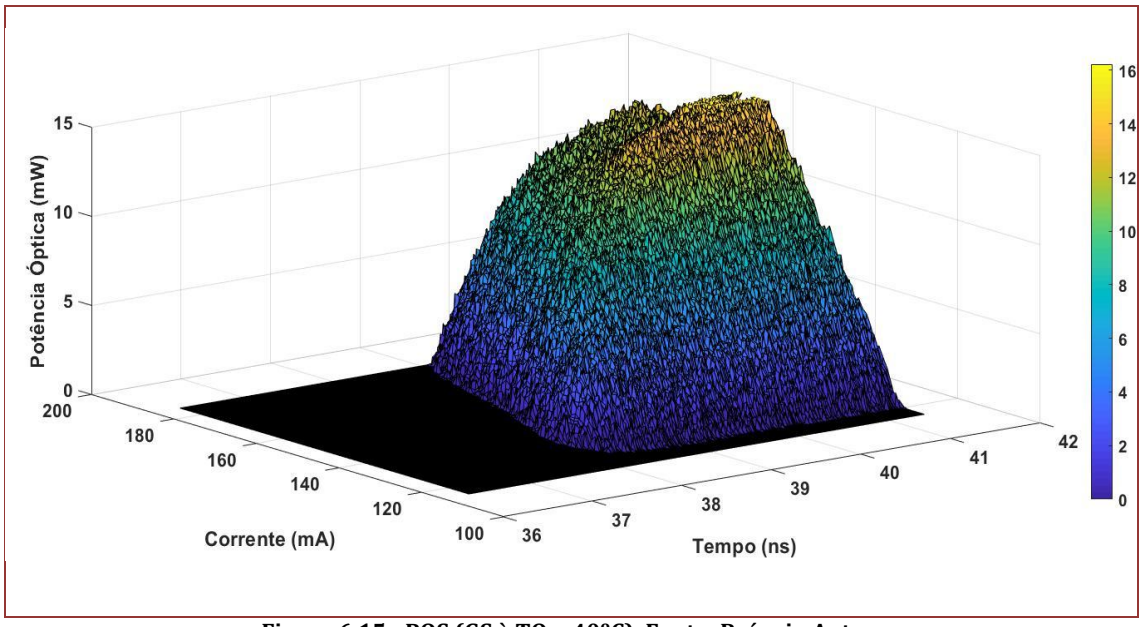


Figura 6.15 - POS (GS à TO = 40ºC). Fonte: Próprio Autor.

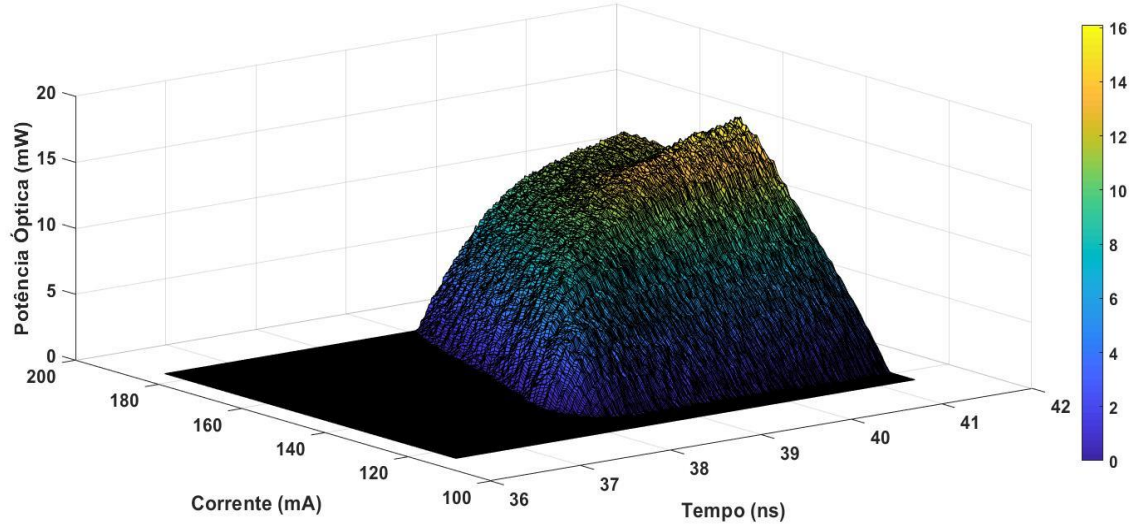


Figura 6.16 – $\widehat{POS}$ do Modelo NARX-ESN-ELM Especialista (GS à TO = 40ºC). Fonte: Próprio Autor.

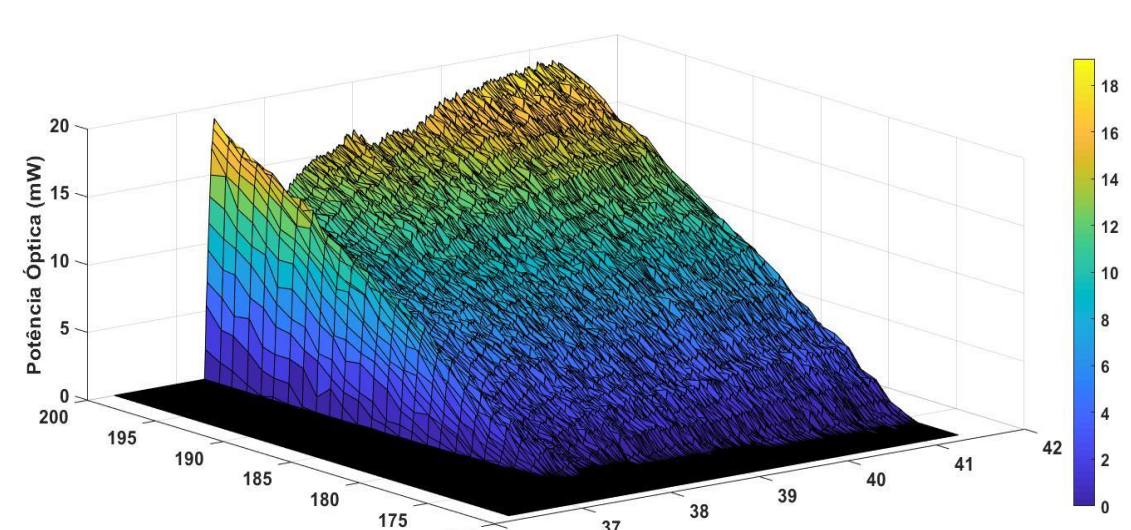


Figura 6.17 - POS (ES à TO = 50ºC). Fonte: Próprio Autor.

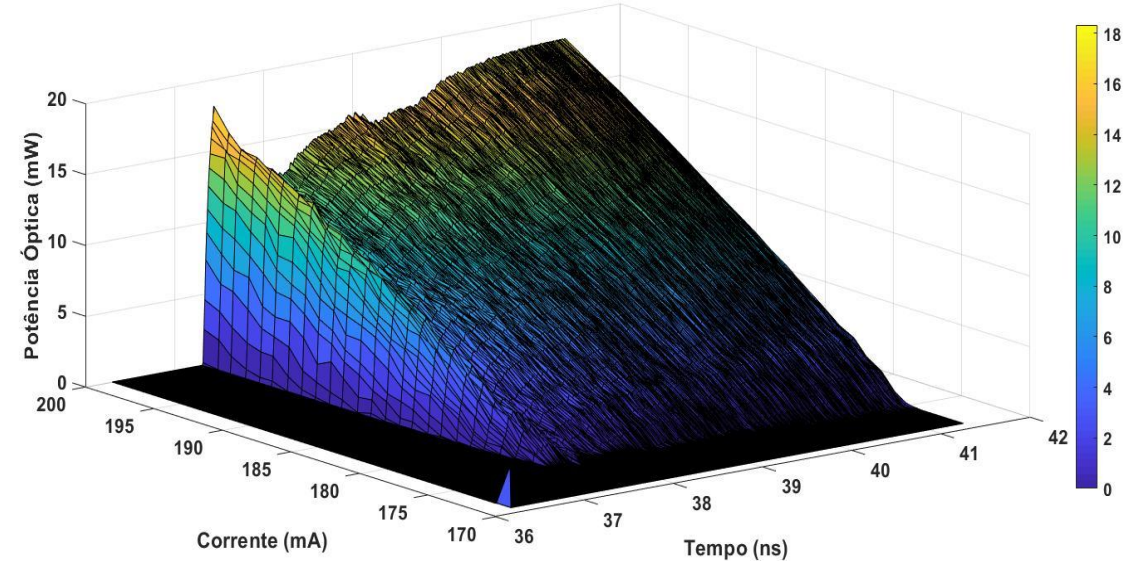


Figura 6.18 - $\widehat{POS}$ do Modelo NARX-ESN-ELM Especialista (ES à TO = 50ºC). Fonte: Próprio Autor.

6.7.ANÁLISE DOS RESULTADOS ESPECÍFICOS

Esta seção possui o objetivo de demonstrar e analisar os resultados graficamente do melhor modelo proposto considerado, o NARX-ESN-ELM de 1º Ordem Especialista, apresentando através das diversas ilustrações o comportamento do LPQ ao longo do tempo com determinadas correntes de injeção, e as curvas P-I.

Tabela 6.4 - Resultados Globais dos Melhores Resultados.

Modelo	RMSE (mW)	MAPE (%)	U de Theil (adimensional)	Tempo Total de Processamento (min)
ELM GENERALISTA	2,30	12,92	0,1156	0,97
ESN GENERALISTA	1,70	9,23	0,0852	2,81
MLP GENERALISTA	0,33	2,15	0,0161	21,60
ELM ESPECIALISTA	2,22	10,98	0,1115	0,43
ESN ESPECIALISTA	15,72	8,83	0,0784	3,10
ABC+ESN ESPECIALISTA	1,97	9,11	0,0982	21.887,99
PSO+ESN ESPECIALISTA	50,37	186,49	0,8413	9.520,40
MLP ESPECIALISTA	0,28	1,51	0,0139	76,30
NARX-ELM ESPECIALISTA	2,70	15,09	0,1365	1,55
NARX-ESN ESPECIALISTA	1,19	15,29	0,0593	0,04
NARX-ESN-ELM ESPECIALISTA	**0,37**	**3,79**	**0,0181**	**0,09**
NARX-MLP ESPECIALISTA	0,29	2,09	0,0141	161,25

6.8.Análise Gráfica das Predições das Séries Temporais

A Figura 6.19 ilustra a potência óptica de saída do LPQ com a TO de 20ºC, tanto no ES quanto no GS, com as correntes de injeção de 102 mA e 140 mA, demonstrando o excelente resultado das simulações com o NARX-ESN-ELM. Tanto na primeira Figura 6.19(a), onde o GS se comporta similarmente a um sistema de 2ºordem, além do ES nulo nesta temperatura de operação, quanto na Figura 6.19(b), onde a saída no ES assemelha-se a um impulso e o GS a um sistema de primeira ordem.

Assim, a Figura 6.19 ilustra todos os comportamentos de saída apresentados nos experimentos, bem como a interação entre as saídas dos estados, principalmente na Figura 6.19(b), onde visualiza-se o decréscimo da saída do ES e paralelamente o surgimento da saída no GS.

Tabela 6.5 - Configurações dos Parâmetros dos Melhores Resultados.

Modelo	Número de Neurônios	Função de inicialização/treinamento	Ordem do NARX	População/Interações ABC/PSO
ELM GENERALISTA	552	Senoidal	-	-
ESN GENERALISTA	1.000	Senoidal	-	-
MLP GENERALISTA	91	Levenberg-Marquardt	-	-
ELM ESPECIALISTA	RNA1: 298; RNA2: 165; RNA3: 242; RNA4:604; RNA5:239	Senoidal	-	
ESN ESPECIALISTA	RNA1: 1.000; RNA2: 1.000; RNA3: 1.000; RNA4: 1.000; RNA5: 1.000.	Senoidal	-	

Tabela 6.6 - Configurações dos Parâmetros dos Melhores Resultados.

Modelo	Número de Neurônios	Função de inicialização/treinamento	Ordem do NARX	População/Interações ABC/PSO
ABC+ESN ESPECIALISTA	RNA1: 100; RNA2: 100; RNA3: 100; RNA4: 100; RNA5: 100;	Senoidal	-	30/50
PSO+ESN ESPECIALISTA	RNA1: 100; RNA2: 100; RNA3: 100; RNA4: 100; RNA5: 100;	Senoidal	-	50/100
MLP ESPECIALISTA	RNA1: 45; RNA2: 34; RNA3: 40; RNA4:47; RNA5:32	Levenberg-Marquardt	-	-
NARX-ELM ESPECIALISTA	RNA1: 820; RNA2: 938; RNA3: 559; RNA4: 511; RNA5: 859	Senoidal	1	-
NARX-ESN ESPECIALISTA	RNA1: 8; RNA2: 9; RNA3: 6; RNA4:31; RNA5: 8.	Senoidal	1	-
NARX-ESN-ELM ESPECIALISTA	**RNA1: 28; RNA2: 58; RNA3: 63; RNA4: 55; RNA5: 98.**	**Senoidal**	**1**	**-**
NARX-MLP ESPECIALISTA	RNA1: 45; RNA2: 34; RNA3: 40; RNA4: 47; RNA5: 32.	Levenberg-Marquardt	4	-

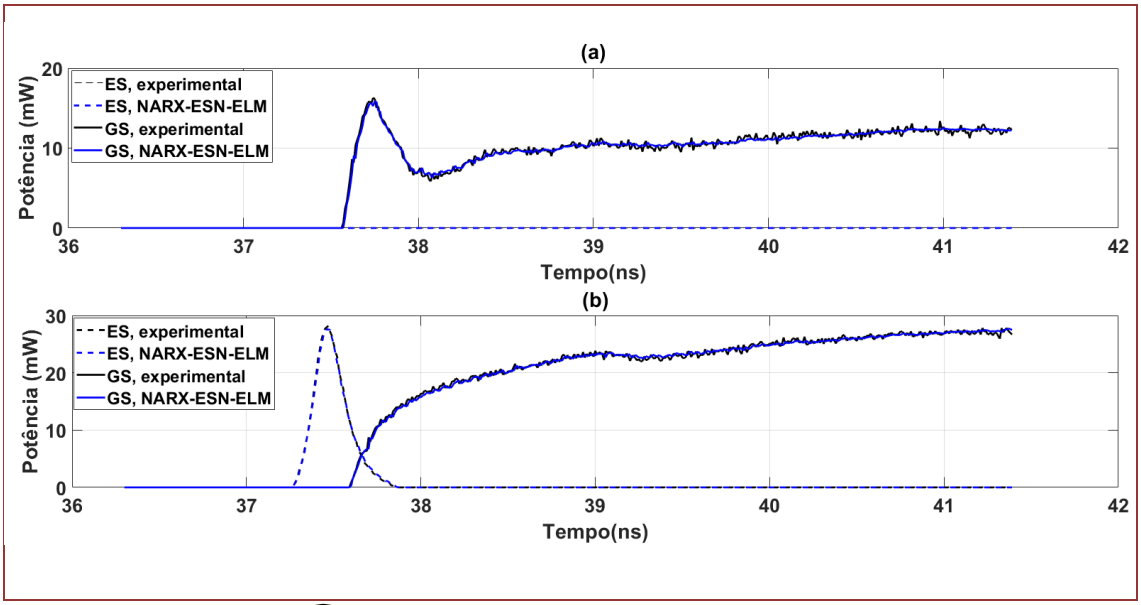


Figura 6.19 – POS x $\widehat{POS}$ à TO = 20ºC. (a) I = 102 mA (b) I = 140 mA. Fonte: Próprio Autor.

Já a Figura 6.20 ilustra a potência óptica de saída do LPQ com a TO de 40ºC, em ambos os estados, com as correntes de injeção de 148 mA e 182 mA, porém, diferentemente do comportamento do LPQ á TO de 20ºC, ambos estados apresentam saída óptica em todo a faixa dinâmica de correntes de injeção analisados nos experimentos, além, da mudança do comportamento do GS para um sistema de 1º ordem e o ES com um sistema de 2ºordem em todas as suas saídas nesta temperatura, e o ES.

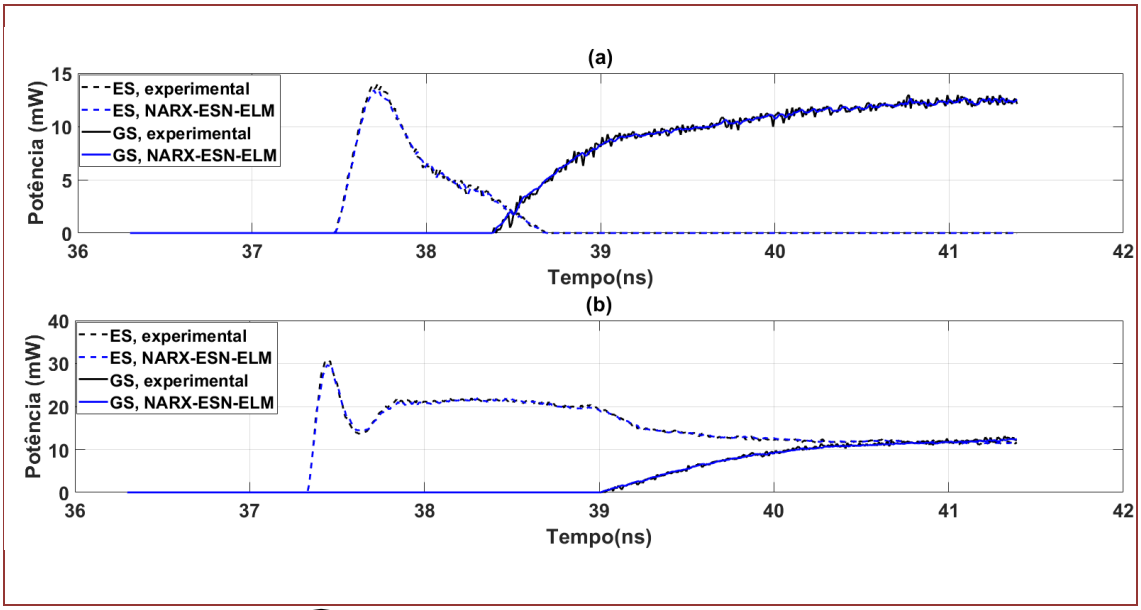


Figura 6.20 - POS x $\widehat{POS}$ à TO = 40ºC. (a) I = 148 mA (b) I = 182 mA. Fonte: Próprio Autor.

Mesmo com esta alteração de comportamento entre as temperaturas, por se tratar de um modelo especialista, o NARX-ESN-ELM conseguiu prever adequadamente as diferentes dinâmicas, já que não há influência entre elas.

Na Figura 6.21 visualiza-se o comportamento da saída óptica do ES à TO de 50ºC, com a ausência do GS, nulo nesta faixa de temperatura, e similar às outras curvas de saída, o modelo escolhido também obteve um ótimo resultado.

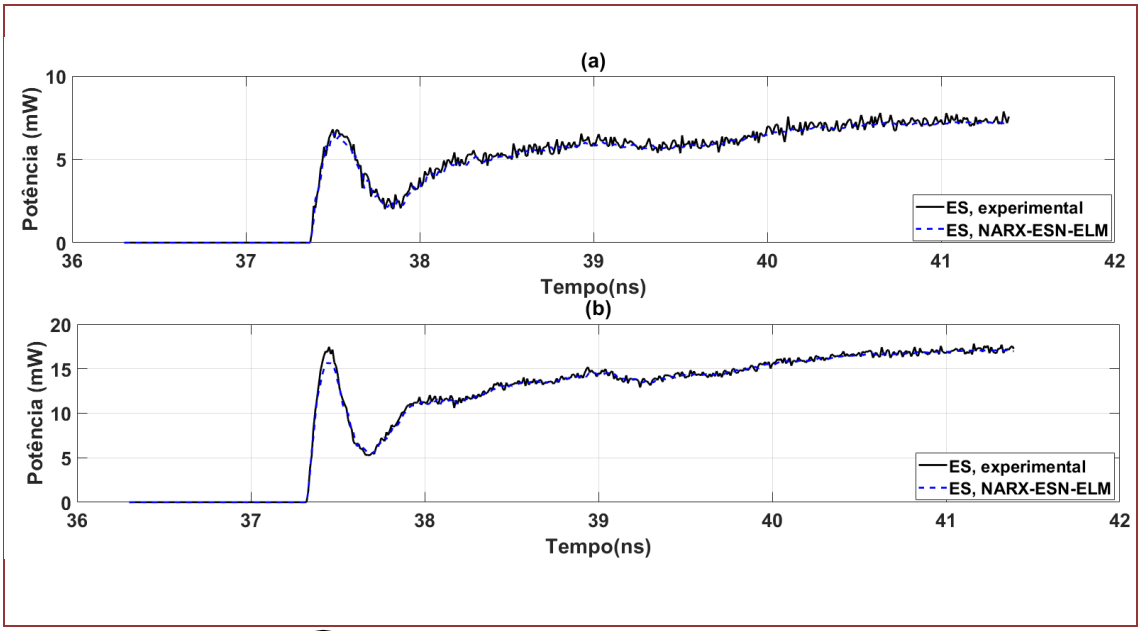


Figura 6.21 – POS x $\widehat{POS}$ à TO = 50ºC. (a) I = 184 mA (b) I = 198 mA. Fonte: Próprio Autor.

6.9.Análise Gráfica do Erro por Séries Temporais

A Figura 6.22 ilustra o RMSE por corrente de injeção nos diversos estados e temperaturas de operação avaliados no trabalho.

Na Figura 6.22(a) é ilustrado o comportamento do RMSE em ambos estados à TO de 20ºC, com o RMSE do GS superior ao ES, consequência do comportamento do LPQ nesta temperatura. Para esta faixa de temperatura, apenas em uma faixa de corrente há POS no ES, diferentemente do GS, que possui em quase a totalidade das correntes de injeções aplicadas nos experimentos.

Também é possível visualizar uma certa uniformidade do RMSE no GS, que se repete na Figura 6.22(b), ou seja, mostrando certa confiabilidade aos resultados obtidos para este estado, independente da temperatura de operação.

A Figura 6.22(b) ilustra o crescimento do RMSE com o aumento da corrente de injeção no ES, demonstrando uma certa dificuldade do modelo proposto em apresentar a mesma confiabilidade do GS. Esta dificuldade é resultante da variação do comportamento especifico do ES e por possuir até a corrente de injeção de 135 mA uma POS nula.

Entretanto, a Figura 6.22(c) ilustra o comportamento do RMSE do ES similar ao GS nas demais temperaturas de operação, devido à similaridade de seus comportamentos.

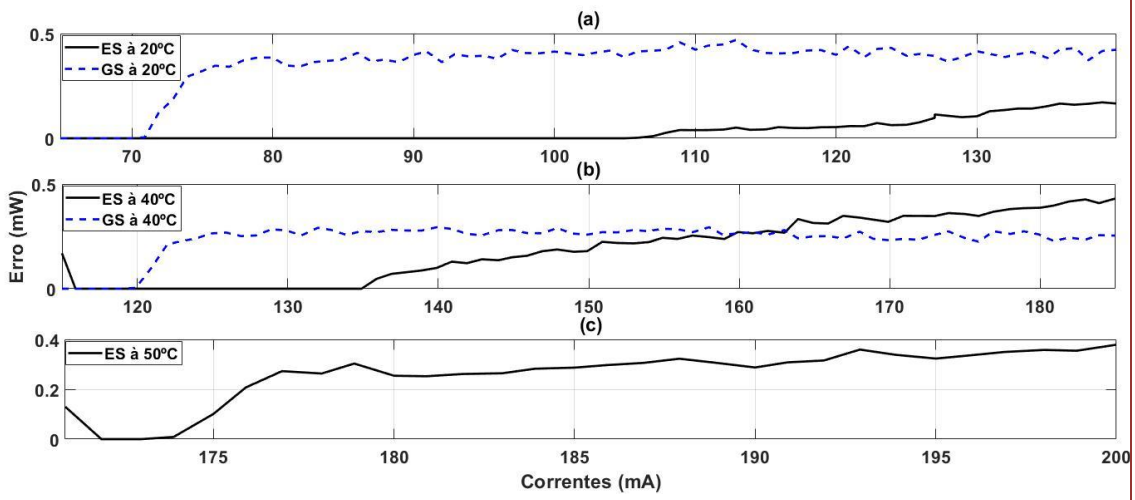


Figura 6.22 – RMSE x Corrente x Estado (a) TO = 20ºC (b) TO = 40ºC e (c) TO = 50ºC. Fonte: Próprio Autor.

As Figuras 6.23 até 6.27 ilustram o gráfico de *Candlestick* do erro para cada par estado/temperatura de operação, proporcionando a análise mais profunda do erro de cada curva, já que destaca não somente a medianas, 1º e 3º quartil dos erros, mas também destaca os valores máximos, mínimos e valores discrepantes, valores não visíveis nos gráficos anteriores.

Assim, a Figura 6.23 apresenta o *Candlestick* das curvas de POS no ES à TO de 20ºC. É possível perceber que após o início da emissão de potência óptica (I > 105 mA), apesar do RMSE ser relativamente baixo (Figura 6.22(a)), apareceu uma grande quantidade de valores discrepantes, em resumo, observou-se, uma piora na predição nas correntes mais elevadas.

Também similar ao ilustrado na Figura 6.22(a), a Figura 6.24 apresenta no *Candlestick* do erro de GS à 20ºC as medianas, 1º e 3º quartil, mínimo e máximo bem

próximos em todas as curvas, e com uma quantidade relativamente baixa de valores discrepantes, já que são 510 amostras por curva, comprovando a aproximada linearidade do erro nas curvas neste caso.

Devido à alteração do comportamento da LPQ no ES à 40ºC em relação ao mesmo estado à 20ºC, a Figura 6.25 apresenta uma menor quantidade de valores discrepantes nas correntes mais elevadas, onde a POS se comporta de maneira mais estável. Já conseguindo-se visualizar as demais informações, como a mediana, além de demonstrar certa estabilidade, diferentemente da primeira metade do gráfico onde o comportamento está em transição.

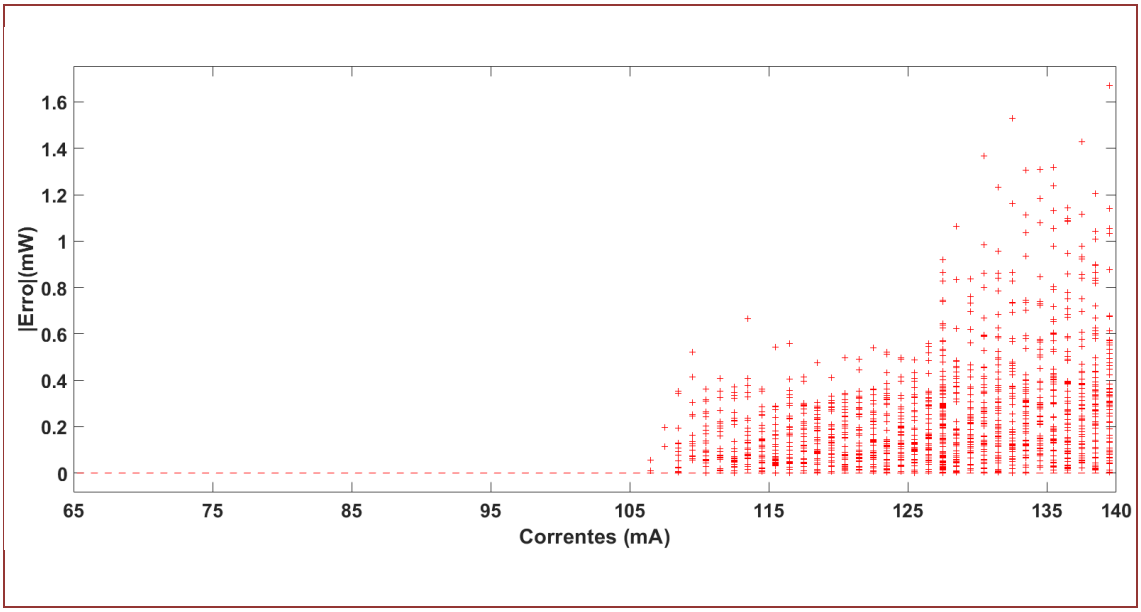


Figura 6.23 - *Candlestick* do LPQ no ES à 20ºC. Fonte: Próprio Autor.

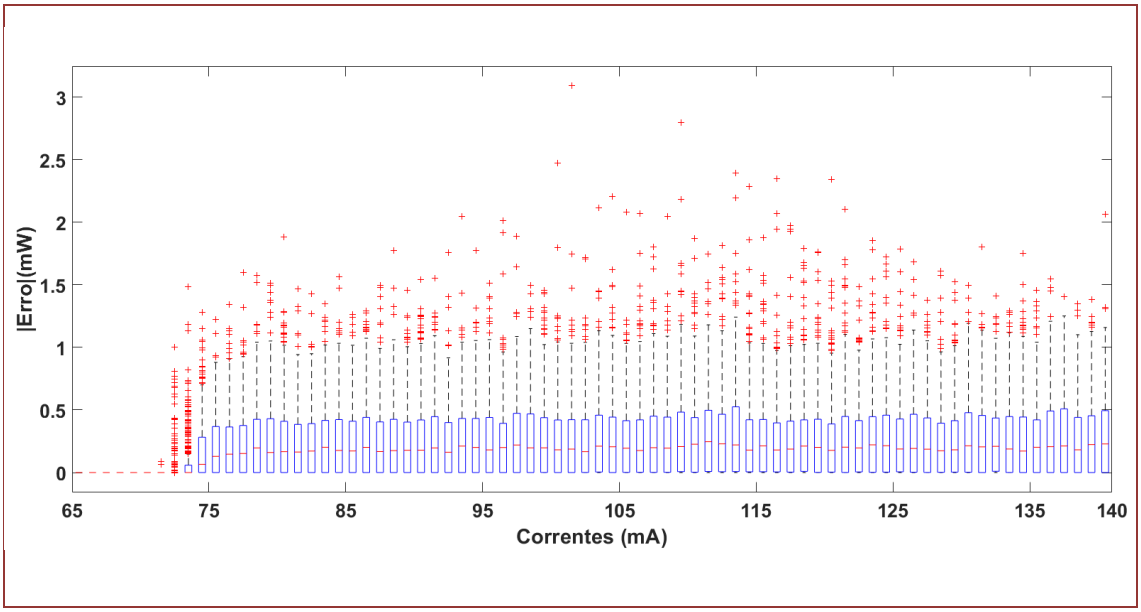


Figura 6.24 - *Candlestick* do LPQ no GS à 20ºC. Fonte: Próprio Autor.

A Figura 6.26 ilustra o *Candlestick* do erro do LPQ no GS à 40ºC, com medianas similares, inclusive diminuindo com a corrente de injeção, e com o incremento da quantidade de valores discrepantes, devido a alteração da curva POS, com um tempo de acomodação superior do que nas correntes de injeção inferiores.

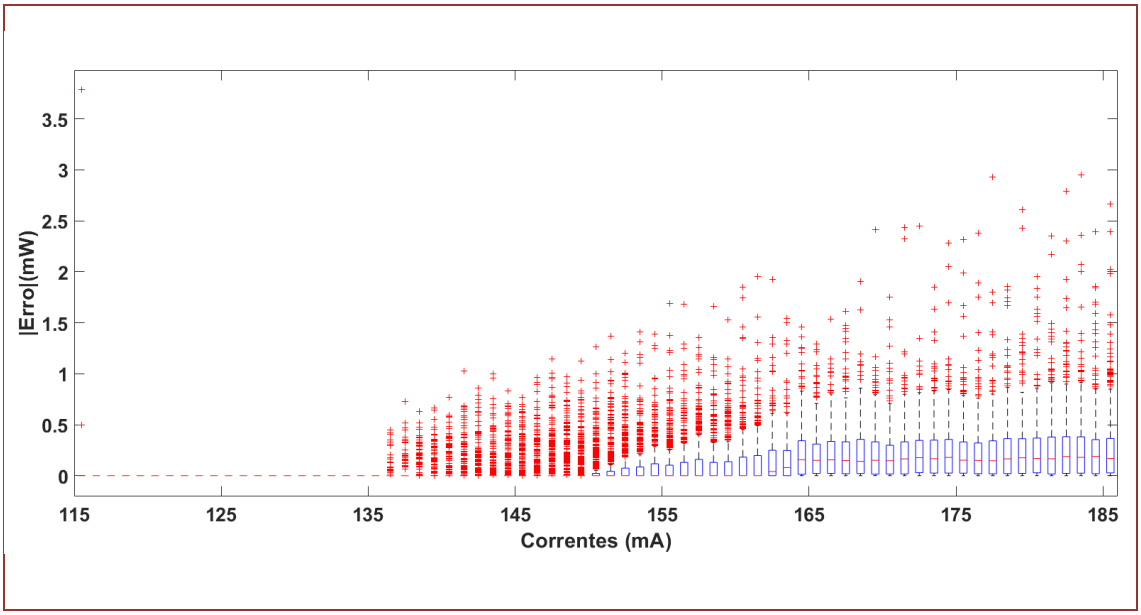


Figura 6.25 - *Candlestick* do LPQ no ES à 40ºC. Fonte: Próprio Autor.

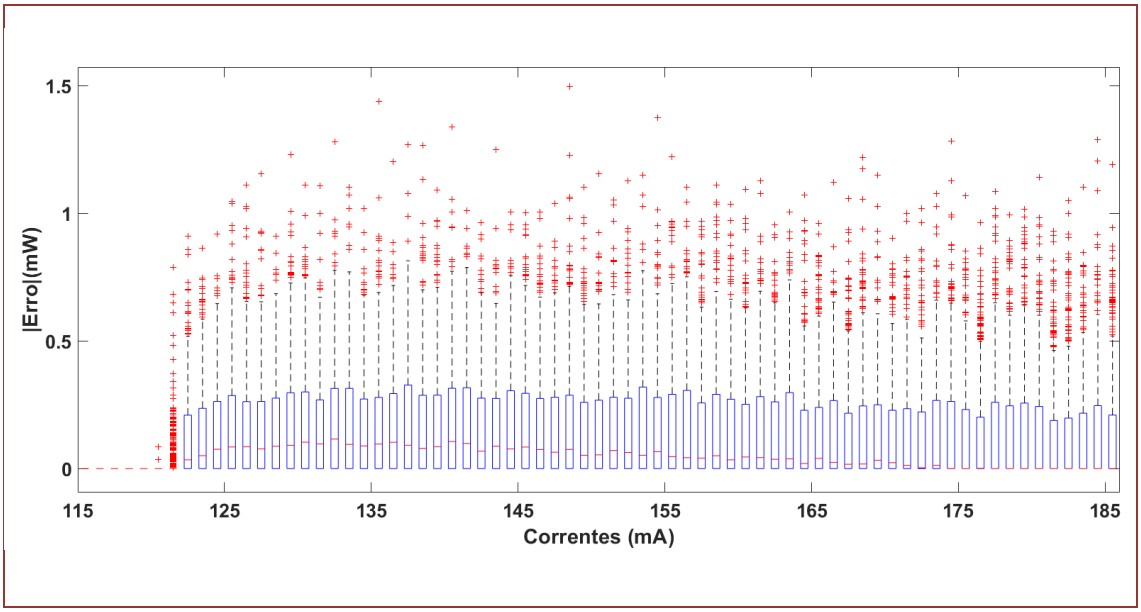


Figura 6.26 - *Candlestick* do LPQ no GS à 40ºC. Fonte: Próprio Autor.

Na Figura 6.27 é apresentado o *Candlestick* do erro do ES à 50ºC, que como seu comportamento se assemelha a um sistema de 2º ordem, seu erro apresenta-se mais estável, com medianas similares entres as diferentes curvas e com uma quantidade de valores discrepantes significativamente inferior as apresentadas nas demais temperaturas de operação, mostrando assim uma predição similar ao GS.

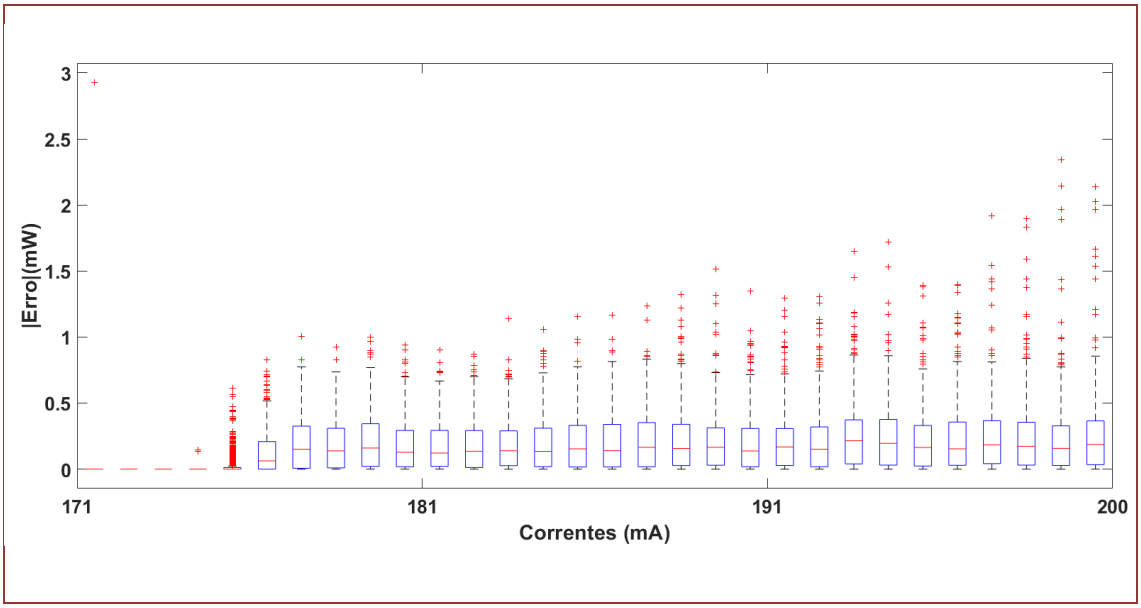


Figura 6.27 - *Candlestick* do LPQ no ES à 50ºC. Fonte: Próprio Autor.

6.10. Curvas P-I

Outra forma muito popular em estudos para avaliação gráfica dos resultados das modelagens da dinâmica do LPQ é a utilização de curvas P-I (Potência x Corrente de injeção na região ativa), ilustradas na Figura 6.28, em todos os estados e temperaturas de operação. As curvas P-I são indispensáveis na avaliação de modelagens, pois permite determinar a corrente limiar do dispositivo a ser projetado, parâmetro muito importante, uma vez que é estritamente relacionado com o consumo de energia do laser (THÉ, 2009).

Outra característica importante do laser que pode ser extraído da curva P-I é a eficiência da inclinação, $\frac{dP}{dI}$, usado para quantificar a energia emitida observada quando a variação da corrente injetada é dI, acima do limite. Este parâmetro estima a eficiência de energia do dispositivo, uma vez maior inclinação significa uma

capacidade de emitir mais energia como resposta a níveis mais baixos de injeção de corrente (THÉ, 2009).

Assim, na Figura 6.28 (a) é apresentada a curva P-I com o LPQ à 20ºC, demonstrando o comportamento similar a uma curva linear no caso do GS e uma curva exponencial no ES. Entretanto, na Figura 6.28 (b), verifica-se uma alteração no comportamento do GS, apresentando um pico em torno da corrente de injeção de 158 mA, em um comportamento similar a um sistema de 2º ordem, enquanto que o ES, tanto na Figura 6.28 (b), quanto na Figura 6.28 (c), continua com seu comportamento similar ao linear.

Em todos os casos da Figura 6.28, é possível verificar a ótima predição pelo NARX-ESN-ELM, podendo assim ser aplicado sem nenhuma restrição na determinação destes parâmetros de produção de dispositivos LPQs.

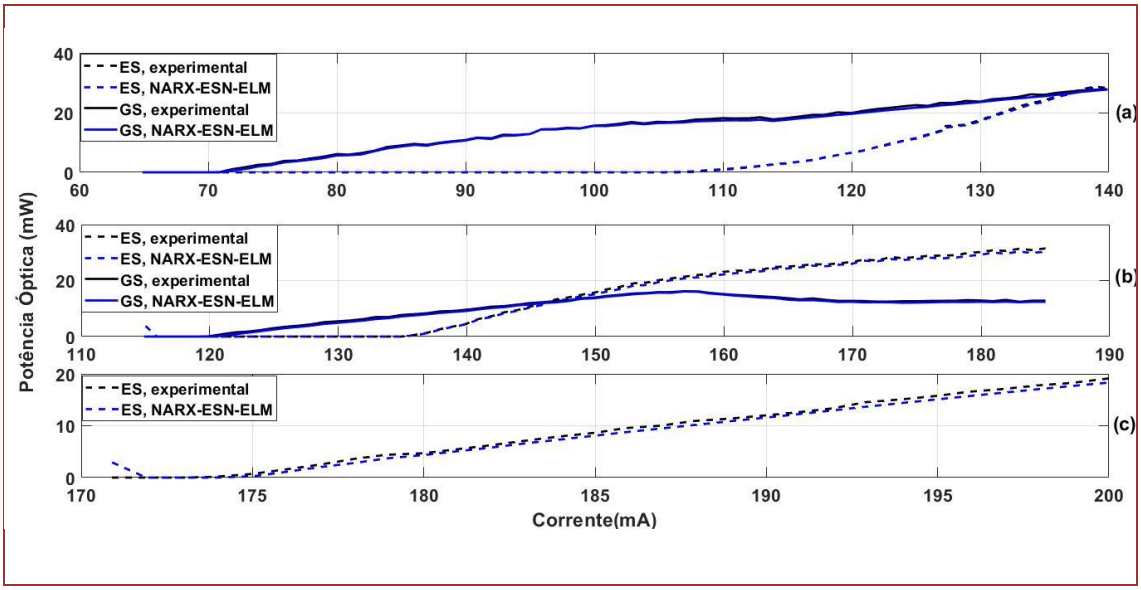


Figura 6.28 – Curva PI (a) TO = 20ºC (b) TO = 40ºC e (c) TO = 50ºC. Fonte: Próprio Autor.

6.11. CONSIDERAÇÕES

Este capítulo apresentou os melhores resultados globais obtidos nos diversos modelos propostos no trabalho, assim como mostrou a análise dos resultados específicos, um detalhamento das curvas de POS e do erro no modelo escolhido, NARX-ESN-ELM de 1ºordem, através de ilustrações gráficas.

7 Conclusões

Atualmente é importante a obtenção de modelos de predição da dinâmica de lasers de pontos quânticos simples, com bom desempenho e baixo tempo total de processamento. Esses novos modelos podem se tornar alternativas para modelos já propostos anteriormente e mais tradicionais, como o modelo baseado em descrição de microestados (GRUNDMANN, HEITZ, *et al.*, 1997) e o modelo baseado em equações de taxa (SUGAWARA, HATORI, *et al.*, 2005).

Neste contexto, este trabalho teve como objetivo desenvolver uma ferramenta de software para a predição da dinâmica de lasers de pontos quânticos com uma capacidade de um treinamento rápido a partir dos dados experimentais, e gerar resultados eficientes à um baixo tempo total de processamento, podendo assim popularizar ainda mais o estudo desta tecnologia.

Assim, foram propostos modelos utilizando RNAs como MLP, ELM e ESN, além de versões híbridas, como a ABC-ESN, PSO-ESN e ESN-ELM. Foi considerado também a utilização do modelo NARX para aprimorar os resultados obtidos.

Em uma primeira análise foram testados os modelos generalistas, onde se utilizou toda a gama de dados experimentais obtidos na modelagem, o que acarretou em uma dificuldade na predição por partes das RNAs, especialmente a ESN (RMSE = 1,70 mW, MAPE = 9,23% e U = 0,0852) e a ELM (RMSE = 2,30 mW, MAPE = 12,92% e U = 0,1156), fazendo-se ser descartada.

Em um segundo instante, confirmou-se a viabilidade da predição da dinâmica em modelos especialistas, separando os dados experimentais pelos pares estado/temperatura de operação, observando-se um incremento no desempenho nas RNAs utilizadas. O MLP (RMSE = 0,28 mW, MAPE = 1,51% e U = 0,0139) obteve seu melhor desempenho, o ELM obteve um RMSE e MAPE inferiores (2,22 mW e 10,98%, respectivamente), enquanto a ESN aumentou o RMSE, apesar de ter diminuído o MAPE (15,72 mA e 8,83 %, respectivamente).

Paralelamente, visando aumentar ainda mais o desempenho da RNA ESN, foram utilizados algoritmos de otimização, o ABC e o PSO, para otimizar a parametrização de suas variáveis de inicialização. Porém, devido à grande gama de valores possíveis e de parâmetros a serem otimizados, os algoritmos não obtiveram êxito nos seus objetivos, sendo descartados, tendo o ABC-ESN um MAPE de 9,11% e o PSO-ESN de 186,49%, além do elevado tempo total de processamento de ambos, com 21,9 mil minutos e 9,5 mil minutos, respectivamente.

Finalmente, em uma última fase, foi utilizado o NARX com as RNAs como estimadores de não linearidade, objetivando a maximização dos resultados, com a menor complexidade possível.

Adicionalmente, foi testado o NARX com a RNA híbrida ESN-ELM, que se apresentou como a melhor solução (RMSE = 0,31 mA, MAPE = 3,79%, U = 0,0181 e C.C. = 0,09 min), para a predição de dinâmica do LPQ, com um tempo total de processamento bastante baixo e o desempenho em relação ao erro satisfatório, equivalente aos encontrados com a RNA MLP (RMSE = 0,28 mW, MAPE = 1,51% e U = 0,0139).

Resumidamente, obteve-se a seguinte avaliação por RNA utilizada:

ELM:

- Tempo baixo;
- Desempenho geral insatisfatório;
- Aplicações: restrita a modelagens onde a saída de interesse seja o valor no regime permanente.

ESN:

- Tempo de processamento extremamente baixo;
- Bom desempenho;
- Aplicações: sem restrições devido ao desempenho.

ESN-ELM:

- Tempo de processamento extremamente baixo;
- Excelente desempenho – ótimo para a predição dos dados;
- Melhor tempo x desempenho;
- Aplicações: sem restrições.

MLP:

- Tempo de processamento elevado;
- Excelente desempenho – ótimo para a predição dos dados;
- Aplicações: sem restrições devido ao desempenho, restrição apenas caso o tempo de processamento seja crucial.

Posteriormente a validação do melhor modelo, foi avaliado mais detalhadamente o desempenho da predição por série temporal, não só para comprovar a viabilidade de seu uso em toda a faixa dinâmica de dados utilizados, mas também para visualização gráfica, e melhor entendimento dos resultados obtidos.

Em suma, conclui-se que a utilização de RNAs, principalmente em NARXs são uma excelente solução e alternativa para a modelagem da dinâmica de lasers semicondutores de pontos quânticos, esperando-se que os resultados obtidos através deste projeto possam ser úteis a comunidade desenvolvedora da tecnologia, tanto no sentido de eficiência na predição, como também forma de popularizar seu desenvolvimento.

Sugestões para trabalhos futuros

Com a experiência adquirida ao longo deste trabalho alguns aspectos de relevância foram observados. Esses podem subsidiar trabalhos futuros, citando-se:

- Comparar resultados obtidos com outras modelagens tradicionais, como modelo baseado em descrição de microestados e o modelo baseado em equações de taxa, visando avaliação tanto do erro quanto do tempo total de processamento;

- Utilizar mais dados experimentais, com diferentes correntes de injeção e temperaturas de operação e testar a eficiência da predição;

- Alterar as entradas da modelagem para as mesmas necessárias em outros modelos, possibilitando não só a comparação de resultados, mas também a viabilidade de transição entre modelos de forma simples;

- Alterar o modelo proposto para poder prever simultaneamente ambos os estados do LPQ, resultando em um modelo de filtragem espectral do LPQ.

Referências

AGUIRRE, L. A. *Introdução à identificação de sistemas:* técnicas lineares e não lineares: teoria e aplicação. 4. ed. Belo Horizonte: UFMG, 2015.

ALAZRAI, R.; LEE, C. S. G. *An NARX-based approach for human emotion identification*. 2012 International Conference on Intelligent Robots and Systems. Vilamoura, Portugal: IEEE/RSJ. 2012. p. 4571-4576.

ALVES, G. A. *Transferência de Energia, Ressonante entre Pontos Quânticos, Zinco Ftalocianina e Praseodímio*. Tese de Doutorado. Uberlândia: Universidade Federal de Uberlândia. 2016.

ASADA, M.; MIYAMOTO, Y.; SUEMATSU, Y. Gain and the threshold of three-dimensional quantum-box lasers. *IEEE Journal of Quantum Electronics*, v. 22, p. 1915-1921, Setembro 1986.

ASRYAN, L. V.; SURIS, R. A. Inhomogeneous line broadening and the threshold current density of a semiconductor quantum dot laser. *Semiconductor Science and Technology*, v. 11, n. 4, p. 554-567, Janeiro 1996.

AUGUSTUS, G. *Inteligência de enxame e o algoritmo das abelhas. Monografia (Graduação em Matemática).* Universidade de São Paulo. São Paulo-SP, p. 27. 2009.

BATTITI, R. First and second order methods for learning: Between steepest descent and Newton's method. *Neural Computation*, v. 4, n. 2, p. 141–166, 1992.

BEN-ISRAEL, A.; THOMAS , T. N. E. *Generalized Inverses*. 2. ed. New York: Springer-Verlag, 2003.

BIMBERG, D. *Quantum dots for lasers, amplifiers and computing*. CLEO/Pacific Rim 2003. The 5th Pacific Rim Conference on Lasers and Electro-Optics. Taipei: IEEE. 2003. p. 2.

BOCCATO, L. *Novas Propostas e Aplicações de Redes Neurais com Estados de Eco. Tese (Doutorado)*. Universidade Estadual de Campinas. Campinas, São Paulo. 2013.

BUSING, L.; SCHRAUWEN, B.; LEGENSTEIN, R. Connectivity, dynamics, and memory in reservoir computing with binary and analog neurons. *Neural Computation*, v. 22, n. 5, p. 1272–1322, Maio 2010.

BUTCHER, J. et al. *Extending reservoir computing with random static projections:* a hybrid between extreme learning and RC. 18th European Symposium on Artificial Neural Networks. Evere, Bélgica: Proceedings. 2010. p. 303-308.

C. SILVA, J. *Confinamento Quânticos em Hetero-estruturas Semicondutoras de Baixa Dimensionalidade. Tese (Doutorado em Física)*. Universidade Federal do Ceará. Fortaleza, p. 162. 2008.

CAO, J.; LIN, Z.; HUANG, G.-B. Self-Adaptive Evolutionary Extreme Learning Machine. *Neural Processing Letters*, v. 36, p. 285–305, Dezembro 2012.

CORCOS, D.; D2, M. *A Procura da Batida Perfeita*. Rio de Janeiro: Sony Music, 2003.

DANTAS, A. D. O. S. *Identificação de modelos polinomiais NARX utilizando algoritmos combinados de detecção de estrutura e estimação de parâmetros com aplicações práticas*. Natal: UFRN, 2013.

DJIE, H. S. et al. Room-temperature broadband emission of an InGaAs/GaAs quantum dots laser. *Optics Letters*, v. 32, n. 1, p. 44-46, 2007.

DRZEWIETZKI, et al. Theoretical and experimental investigations of the temperature dependent continuous wave lasing characteristics and the switch-on dynamics of an InAs/InGaAs quantum-dot semiconductor laser. *Optics Communications*, v. 283, p. 5092-5098, Dezembro 2010.

ENDAH, S. N. et al. *Beyond back-propagation learning for diabetic detection:* Convergence comparison of gradient descent, momentum and Adaptive Learning Rate. 1st International Conference on Informatics and Computational Sciences. Semarang: ICICoS. 2017. p. 189-194.

ERNEUX, T.; GLORIEUX, P. *Laser Dynamics*. Cambridge: Cambridge University Press, 2010.

EZRA, Y. B.; LEMBRIKOV, B. I.; ZARKOVSKY, S. *Applications of quantum dot (QD) lasers in optical communications for datacenters*. 19th International Conference on Transparent Optical Networks. Girona: ICTON. 2017. p. 1-4.

FAN, H. T.; WANG, W.; JIN, Z. *Performance optimization of echo state networks through principal neuron reinforcement*. 2017 International Joint Conference on Neural Networks (IJCNN). Anchorage: IEEE. 2017. p. 1717-1723.

GHASEMI-FARD, M. et al. *A low cost and reliable Polak-Ribiere conjugate gradient deflation ICA algorithm for real signals*. 21st Iranian Conference on Electrical Engineering. Mashhad: [s.n.]. 2013. p. 1-5.

GIOANNINI, M.; MONTROSSET, I. Numerical Analysis of the Frequency Chirp in Quantum-Dot Semiconductor Lasers. *IEEE Journal of Quantum Electronics*, v. 43, n. 10, Outubro 2007.

GOLDBARG, M. C.; LUNA, H. P. L. *Otimização Combinatória e Programação Linear:* Modelos e Algoritmos. 2. ed. [S.l.]: Campus, 2005.

GOLUB, G. H.; VAN LOAN, C. F. *Matrix Computations*. 3. ed. Baltimore: Johns Hopkins, 1996.

GRUNDMANN, M. et al. Carrier Dynamics in Quantum Dots: Modeling with Master Equations for the Transitions between Micro-States. *Physica Status Solidi B*, v. 203, p. 121-132, Novembro 1997.

HAGAN, M. T.; DEMUTH, H. B.; BEALE, M. H. *Neural Network Design*. Boston: PWS Publishing, 1996.

HAYKIN, S. *Neural networks and learning machines*. 3. ed. New Jersey: Pearson Education, 2008.

HOOGLAND, S. The Fuss About Quantum Dots. *Photonics Spectra*, Janeiro 2008.

HUANG, G. et al. Trends in extreme learning machines: A review. *Neural Networks*, v. 61, p. 32-48, Janeiro 2015.

HUANG, G.-B.; CHEN, L. Convex incremental extreme learning machine. *Neurocomputing*, v. 70, p. 3056-3062, Outubro 2007.

HUANG, G.-B.; ZHU, Q.-Y.; SIEW, C.-K. Extreme learning machine: Theory and applications. *Neurocomputing*, v. 70, p. 489–501, Maio 2006.

JAEGER, H. The “echo state” approach to analysing and training recurrent neural networks – with an Erratum note. *Fraunhofer Institute for Autonomous Intelligent Systems*, p. 1-47, Janeiro 2010.

KANTNER, M. et al. *Simulation of quantum dot devices by coupling of quantum master equations and semi-classical transport theory*. 2017 International Conference on Numerical Simulation of Optoelectronic Devices. Copenhagen: NUSOD. 2017. p. 217-218.

KANTNER, M.; MITTNENZWEIG, M.; KOPRUCKI, T. Hybrid quantum-classical modeling of quantum dot devices. *Phys. Rev. B*, n. 96, Novembro 2017.

KARABOGA, D. *An Idea Based on Honey Bee Swarm for Numerical Optimization*. echnical Report-TR06, Department of Computer Engineering, Engineering Faculty. Kayseri: Erciyes University. 2005.

KENNEDY, J.; EBERHART, R. *Particle swarm optimization*. IEEE International Conference on Neural Networks. Perth, Australia: Proceedings. 1995. p. 1942-1948.

KHADSE, C. B.; CHAUDHARI, M. A.; BORGHATE, V. B. Electromagnetic Compatibility Estimator Using Scaled Conjugate Gradient Backpropagation Based Artificial Neural Network. *IEEE Transactions on Industrial Informatics*, v. 13, n. 3, p. 1036-1045, Junho 2017.

KHAFAF, N. A.; EL-HAG, A. Bayesian regularization of neural network to predict leakage current in a salt fog environment. *IEEE Transactions on Dielectrics and Electrical Insulation*, v. 25, n. 2, p. 686-693, Abril 2018.

KRAKOWSKI, M. et al. *High Power, Very Low Noise, C.W. Operation of 1.32μm Quantum-Dot Fabry-Perot Laser Diodes*. 20th International Semiconductor Laser Conference. Kohala: IEEE. 2006. p. 39-40.

LARGER, L.; DUDLEY, J. M. Nonlinear dynamics: Optoelectronic chaos. *Nature*, v. 465, p. 41–42, Maio 2010.

LATHI, B. P. *Sinais e Sistemas Lineares*. 2. ed. [S.l.]: Bookman, 2007.

LENKA, S. K.; MOHAPATRA, A. G. *Gradient Descent with Momentum based Neural Network Pattern Classification for the Prediction of Soil Moisture Content in Precision Agriculture*. 2015 IEEE International Symposium on Nanoelectronic and Information Systems. Indore: [s.n.]. 2015. p. 63-66.

LIN, Z. et al. *Modeling and Simulation of the Multi-Population Quantum Dot Lasers Based on Equivalent Circuit*. XXV Physics and Simulation of Optoelectronic Devices. São Francisco: [s.n.]. Fevereiro 2017.

LINGNAU, B. *Nonlinear and Nonequilibrium Dynamics of Quantum-Dot Optoelectronic Devices*. Berlin: Springer, 2015.

LÜDGE, K. *Nonlinear Laser Dynamics:* From Quantum Dots to Cryptography. Weinheim: Wiley-VCH, 2012.

LUTHRA, I. et al. *Comparative study on nature inspired algorithms for optimization problem*. International Conference on Electronics, Communication and Aerospace Technology. Coimbatore: ICECA 2017. 2017. p. 143-147.

MAASS, W.; NATSCHLÄGER, T.; MARKRAM, H. Real-time computing without stable states: a new framework for neural computation based on perturbations. *Neural Computation*, v. 14, n. 11, p. 2531-60, Novembro 2002.

MENEZES JÚNIOR, J. M. P. *Contribuições ao Problema de Predição Recursiva de Séries Temporais Univariadas Usando Redes Neurais Recorrentes. Tese (Doutorado em Engenharia de Teleinformática).* Universidade Federal do Ceará. Fortaleza, p. 188. 2012.

MORETTIN, P. A.; TOLOI, C. M. C. *Análise de séries temporais*. 2. ed. São Paulo: Edgard Blucher, 2006.

NELLES, O. *Nonlinear system identification:* from classical approaches to neural networks and fuzzy models. New York: Springer Science & Business Media, 2013.

NISHI, K. et al. Development of Quantum Dot Lasers for Data-Com and Silicon Photonics Applications. *IEEE Journal of Selected Topics in Quantum Electronics*, v. 23, n. 6, p. 1-7, Novembro 2017.

OGATA, K. *Engenharia de Controle Moderno*. 5. ed. Boston: Pearson, 2010.

PACÍFICO, L. D.; LUDEMIR, T. B. *Melhorando Redes Neurais do Tipo Extreme Learning Machine Através da Otimização por Enxame de Partículas com Mecanismo de Seleção*. XXI Brazilian Symposium on Artificial Intelligence. Curitiba: BRACIS. 2012. p. 1-11.

PAGNOSSIN, I. R. *Pontos-quânticos: fotodetectores, localização-fraca e estados de borda contra-rotativos. Tese (Doutorado em Ciências)*. Instituto de Física. São Paulo, p. 106. 2008.

POWELL, M. J. D. Restart procedures for the conjugate gradient method. *Mathematical Programming*, v. 12, p. 241–254, 1977.

QIAO, J. et al. Growing Echo-State Network With Multiple Subreservoirs. *IEEE Transactions on Neural Networks and Learning Systems*, v. 28, n. 2, p. 391-404, Fevereiro 2017.

RAFAILOV, E. U.; CATALUNA, M. A.; SIBBETT, W. High-power picosecond and femtosecond pulse generation from a two-section mode-locked quantum-dot laser. *Applied Physics Letters*, v. 87, Agosto 2005.

RAFAILOV, E. U.; CATALUNA, M. A.; SIBBETT, W. Mode-locked quantum-dot lasers. *Nature Photonics*, v. 1, p. 395–401, Julho 2007.

RIEDMILLER, M.; BRAUN, H. *A direct adaptive method for faster backpropagation learning:* the RPROP algorithm. IEEE International Conference on Neural Networks. São Francisco: IEEE. 1993. p. 586-591.

SALAH, K. *A Generic Model Order Reduction Technique Based On Particle Swarm Optimization (PSO) Algorithm*. IEEE EUROCON 2017. OHRID: [s.n.]. 2017. p. 4.

SANTOS, W. F. S.; SIMAS FILHO, E. F.; THÉ, G. A. P. *Predição da dinâmica de lasers de pontos quânticos semicondutores utilizando máquinas de aprendizado extremo*. Vitória-ES: XXI Congresso Brasileiro de Automática. 2016.

SAPNA, S.; TAMILARASI, A.; KUMA, M. P. *Backpropagation Learning Algorithm Based on Levenberg Marquardt Algorithm*. The Fourth International Workshop on Computer Networks & Communications. Melbourne: NCOM. 2012.

SCHWEIZER., H.; JETTER, M.; SCHOLZ, F. Quantum-Dot Lasers. *Topics in Applied Physics*, Berlin, v. 90, n. Springer, p. 185-236, Novembro 2003.

SHI, H.-Y.; ZHAO, W.-F.; SU, Y. *Balanced Artificial Bee Colony Algorithm Based on Multiple Selections*. 2017 3rd International Conference on Control, Automation and Robotics. Nagoya: IEEE. 2017.

SILVA , T. G. *Caracterização Óptica e Morfológica de Microestruturas Auto-organizadas e Filmes de Óxido de Zinco (ZnO) Fabricados por Pirólise com Nebulizador Ultrassônico. Dissertação (Mestrado em Física).* Universidade Federal do Paraná - UFPR. Curitiba, p. 118. 2010.

SILVA, M. J. *Crescimento e caracterização de ponto-quânticos de inas auto-formados. Tese (Mestrado em Física)*. Universidade de São Paulo (USP). São Paulo. 1999.

SIQUEIRA, H. V. *Máquinas Desorganizadas para Previsão de Séries de Vazões. Tese (Doutorado em Engenharia Elétrica).* Universidade Estadual de Campinas. Campinas, p. 244. 2013.

SORIANO, M. C. et al. Complex photonics: Dynamics and applications of delay-coupled semiconductors lasers. *Reviews of Modern Physics*, v. 85, p. 421-470, 2013.

SUGAWARA, M. et al. Effect of homogeneous broadening of optical gain on lasing spectra in self-assembled Inx Ga1-x As/GaAs quantum dot lasers. *Physical Review B*, v. 61, p. 7595, Março 2000.

SUGAWARA, M. et al. Modeling room-temperature lasing spectra of 1.3-μm self-assembled InAs/GaAs quantum-dot lasers: Homogeneous broadening of optical gain under current injection. *Journal of Applied Physics*, v. 97, Janeiro 2005.

SUMMERS, D. et al. The Influence of the Gain—Carrier Density Characteristic on Q-Switching in Quantum-Dot Lasers. *IEEE Journal of Selected Topics in Quantum Electronics*, v. 13, n. 5, p. 1222-1226, Setembro 2007.

THÉ, G. A. P. How to simulate a semiconductor quantum dot laser: general description. *Revista Brasileira de Ensino de Física*, v. 31, p. 2302.1-2302.8, 2009.

THÉ, G. A. P. *Improved modeling and simulation of quantum dot laser*. Tese de Doutorado. Turin: Politecnico di Torino. 2010.

THEIL, H. Applied Economic Forecasting. *The Economic Journal*, Amsterdam, 1966.

VESELINOV, K. et al. Modelling of the two-state lasing and the turn-on delay in 1.55 μm InAs/InP (113)B quantum dot lasers. *IEE Proceedings - Optoelectronics*, v. 153, n. 6, p. 308-311, Dezembro 2006.

YANG, C.; ZHU, X.; WANG, L. *A Novel Reservoir Design Strategy for Echo State*. 2017 Chinese Automation Congress (CAC). Jinan: IEEE. 2017. p. 1484-1488.

YAVARI, M. H.; AHMADI, V. Circuit-Level Implementation of Semiconductor Self-Assembled Quantum Dot Laser. *IEEE Journal of Selected Topics in Quantum Electronics*, v. 15, n. 3, p. 774-779, Maio 2009.

ZHANG, G.; PATUWO, B. E.; HU, M. Y. Forecasting with artificial neural networks: The state of the art. *International Journal of Forecasting*, v. 14, n. 1, p. 35-62, 1998.

ZHANG, Y.; MU, B.; ZHENG, H. Link Between and Comparison and Combination of Zhang Neural Network and Quasi-Newton BFGS Method for Time-Varying Quadratic Minimization. *IEEE Transactions on Cybernetics*, v. 43, n. 2, p. 490-503, Abril 2013.

SOBRE O AUTOR

WILD FREITAS DA SILVA SANTOS – Doutorando em Engenharia Elétrica pelo Programa de Pós-Graduação em Engenharia Elétrica da Universidade Federal da Bahia. Mestre em Engenharia Elétrica pelo Programa de Pós-Graduação em Engenharia Elétrica da Universidade Federal da Bahia (2018) na área de Processamento e Transmissão da Informação. Possui especialização em MBA Executivo em Finanças, Controladoria e Auditoria pela Fundação Getúlio Vargas (2017) e em MBA Executivo em Gestão de Projetos pelo SENAI CIMATEC (2014). Possui graduação em Engenharia Elétrica pela Universidade Salvador (2010). Atua nas áreas de pesquisa aprendizado de máquina, redes neurais artificiais e Iluminação Pública.

www.poisson.com.br
contato@poisson.com.br

@editorapoisson

https://www.facebook.com/editorapoisson